

Size versus truthfulness in the House Allocation problem*

Piotr Krysta¹, David Manlove², Baharak Rastegari² and Jinshan Zhang¹

¹ *Department of Computer Science, University of Liverpool,
Ashton Building, Ashton Street, Liverpool L69 3BX, UK.
Email: {p.krysta, jinshan.zhang}@liverpool.ac.uk.*

² *School of Computing Science, University of Glasgow,
Sir Alwyn Williams Building, Glasgow G12 8QQ, Glasgow, UK.
Email: {david.manlove, baharak.rastegari}@glasgow.ac.uk.*

Abstract

We study the House Allocation problem (also known as the Assignment problem), i.e., the problem of allocating a set of objects among a set of agents, where each agent has ordinal preferences (possibly involving ties) over a subset of the objects. We focus on truthful mechanisms without monetary transfers for finding large Pareto optimal matchings. It is straightforward to show that no deterministic truthful mechanism can approximate a maximum cardinality Pareto optimal matching with ratio better than 2. We thus consider randomized mechanisms. We give a natural and explicit extension of the classical Random Serial Dictatorship Mechanism (RSDM) specifically for the House Allocation problem where preference lists can include ties. We thus obtain a universally truthful randomized mechanism for finding a Pareto optimal matching and show that it achieves an approximation ratio of $\frac{e}{e-1}$. The same bound holds even when agents have priorities (weights) and our goal is to find a maximum weight (as opposed to maximum cardinality) Pareto optimal matching. On the other hand we give a lower bound of $\frac{18}{13}$ on the approximation ratio of any universally truthful Pareto optimal mechanism in settings with strict preferences. In the case that the mechanism must additionally be non-bossy, an improved lower bound of $\frac{e}{e-1}$ holds. This lower bound is tight given that RSDM for strict preference lists is non-bossy. We moreover interpret our problem in terms of the classical secretary problem and prove that our mechanism provides the best randomized strategy of the administrator who interviews the applicants.

Keywords: House allocation problem; Assignment problem; Pareto optimal matching; Randomized mechanisms; Truthfulness

1 Introduction

We study the problem of allocating a set of indivisible objects among a set of agents. Each agent has private ordinal preferences over a subset of objects — those they find acceptable, and each agent may be allocated at most one object. This problem has been studied by both economists and computer scientists. When monetary transfers are not permitted,

*Supported by grants EP/K0100X/1 and EP/K010042/1 from the Engineering and Physical Sciences Research Council.

the problem is referred to as the *House Allocation problem* (henceforth abbreviated by HA; see, e.g., [19, 1, 34]) or the *Assignment problem* (see, e.g., [18, 9]) in the literature. In this paper we opt for the term House Allocation problem. Most of the work in the literature assumes that the agents have strict preferences over their acceptable objects. It often happens though that an agent is indifferent between two or more objects. Here we let agents express indifference, and hence preferences may involve ties unless explicitly stated otherwise.

It is often desired that as many objects as possible become allocated among the agents — i.e., an allocation of maximum size is picked, hence making as many agents happy as possible. Usually, depending on the application of the problem, we are required to consider some other optimality criteria, sometimes instead of and sometimes in addition to maximizing the size of the allocation. Several optimality criteria have been considered in the HA setting, and perhaps the most studied such concept is *Pareto optimality* (see, e.g., [1, 2, 14, 12, 29]), sometimes referred to as Pareto efficiency. Economists, in particular, regard Pareto optimality as the most fundamental requirement for any “reasonable” solution to a non-cooperative game. Roughly speaking, an allocation μ is Pareto optimal if there does not exist another allocation μ' in which no agent is worse off, and at least one agent is better off, in μ' . In this work we are mainly concerned with Pareto optimal allocations of maximum size, but will also consider weighted generalisations.

The related *Housing Market* problem (HM) is the variant of HA in which there is an *initial endowment*, i.e., each agent owns a unique object initially (in this case the numbers of agents and objects are usually defined to be equal). In this setting, the most widely studied solution concept is that of the *core*, which is an allocation of agents to objects satisfying the property that no coalition C of agents can improve (i.e., every agent in C is better off) by exchanging their own resources (i.e., the objects they brought to the market). In the case of strict preferences, the core is always non-empty [27], unique, and indeed Pareto optimal. When preferences may include ties, the notion of *core* that we defined is sometimes referred to as the *weak core*. In this case a core allocation need not be Pareto optimal. Jaramillo and Manjunath [20], Plaxton [24], and Saban and Sethuraman [30] provide polynomial-time algorithms for finding a core allocation that does additionally satisfy Pareto optimality. Our problem differs from HM in that there is no initial endowment, and hence our focus is on Pareto optimal matchings rather than outcomes in the core.

For strictly ordered lists, Abraham et al. [2] gave a characterisation of Pareto optimal matchings that led to an $O(m)$ algorithm for checking an arbitrary matching for Pareto optimality, where m is the total length of the agents’ preference lists. This characterisation was extended to the case that preference lists may include ties by Manlove [22, Sec. 6.2.2.1], also leading to an $O(m)$ algorithm for checking a matching for Pareto optimality. For strictly ordered preference lists, a maximum cardinality Pareto optimal matching can be found in $O(\sqrt{n_1}m)$ time, where n_1 is the number of agents [2]. The fastest algorithm currently known for this problem when preference lists may include ties is based on minimum cost maximum cardinality matching and has complexity $O(\sqrt{nm} \log n)$ (see, e.g., [22, Sec. 6.2.2.1], where n is the total number of agents and objects).

As stated earlier, agents’ preferences are private knowledge. Hence, unless they find it in their own best interests, agents may not reveal their preferences truthfully. An allocation mechanism is *truthful* if it gives agents no incentive to misrepresent their preferences. Perhaps unsurprisingly, a mechanism based on constructing a maximum cardinality Pareto optimal allocation is manipulable by agents misrepresenting their preferences (Theorem 2.1 in Section 2). Hence, we need to make a compromise and weaken at least one of these requirements. In this work, we relax our quest for finding a maximum cardinality Pareto

optimal allocation by trading off the size of a Pareto optimal allocation against truthfulness; more specifically, we seek truthful mechanisms that provide good approximations to the maximum size.

Under strict preferences, Pareto optimal matchings can be computed by a classical algorithm called the *Serial Dictatorship Mechanism* (SDM) — see, e.g., [1], also referred to as the *Priority Mechanism* — see, e.g., [9]. SDM is a straightforward greedy algorithm that takes each agent in turn and allocates him to the most preferred available object on his preference list. Precisely due to this greedy approach, SDM is truthful. Furthermore, SDM is guaranteed to find a Pareto optimal allocation at least half the size of a maximum one, merely because *any* Pareto optimal allocation is at least half the size of a maximum one (see, e.g., [2]). Hence, at least in the case of strict preferences, we are guaranteed an approximation ratio of 2. Can we do better? It turns out that if we stay in the realm of deterministic mechanisms, a 2-approximation is the best we can hope for (Theorem 2.1, Section 2).

Hence we turn to randomized mechanisms in order to achieve a better approximation ratio. The obvious candidate to consider is the Random Serial Dictatorship Mechanism (RSDM) (see, e.g., [1]), also known as the Random Priority mechanism (see, e.g., [9]), that is defined for settings with strict preferences. RSDM randomly generates an ordering of the agents and then proceeds by running SDM relative to this ordering. It is straightforward to extend results from [8] in order to show that RSDM has an approximation ratio of $\frac{e}{e-1}$ (relative to the size of a maximum cardinality Pareto optimal matching).

When indifference is allowed, finding a Pareto optimal allocation is not as straightforward as for strict preferences. For example, one may consider breaking the ties randomly and then applying SDM. This approach, unfortunately, may produce an allocation that is not Pareto optimal (see beginning of Section 3). Few works in the literature considered extensions of SDM to the case where agents' preferences may include ties. However Bogomolnaia and Moulin [10] and Svensson [31] provide an implicit extension of SDM (in the former case for *dichotomous preferences*¹) but do not describe how one may actually implement the algorithm. Aziz et al. [4] provide an explicit extension for a more general class of problems, including HA. Pareto optimal matchings in HA can also be found by reducing to the HM setting [20], which involves creating dummy objects as endowments for the agents. This enables us to utilize one of the aforementioned algorithms [20, 24, 30] to find a Pareto optimal matching in the core. However the reduction increases the instance size, and in particular the number of agents n_1 and the maximum length of a tie in any agent's preference list. Consequently, even the fastest truthful Pareto optimal mechanism for HM, that of Saban and Sethuraman [30], has time complexity no better than $O(n_1^3)$ in the worst case.

Contributions of this paper. In this paper we provide a natural and explicit extension of SDM for the setting in which preferences may exhibit indifference. We argue that our extension is more intuitive than that of Aziz et al. [4] when considering specifically HA. Moreover, as the mechanism of Saban and Sethuraman [30] does not consider the agents sequentially, it is difficult to analyse its approximation guarantee. Our algorithm runs in time $O(n_1^2\gamma)$, where γ is the maximum length of a tie in any agent's preference list. This is faster than the algorithm in [30] when $\gamma < n_1$. We prove the following results that involve upper and lower bounds for the approximation ratio (relative to the size of a maximum cardinality Pareto optimal matching) of randomized truthful mechanisms for computing a Pareto optimal matching:

¹An agent's preference list is *dichotomous* if it comprises a single tie containing all acceptable objects.

1. By extending RSDM to the case of preference lists with ties, we give a universally truthful randomized mechanism² for finding a Pareto optimal matching that has an approximation ratio of $\frac{e}{e-1}$ with respect to the size of a maximum cardinality Pareto optimal matching.
2. We give a lower bound of $\frac{18}{13}$ on the approximation ratio of any universally truthful Pareto optimal mechanism in settings with strict preferences. If the mechanism must additionally be non-bossy³, we observe that [7] implies an improved lower bound of $\frac{e}{e-1}$. This lower bound is tight given that the classical RSDM mechanism for strict preferences is non-bossy.
3. We extend RSDM to the setting where agents have priorities (weights) and our goal is to find a maximum weight (as opposed to maximum cardinality) Pareto optimal matching. Our mechanism is universally truthful and guarantees a $\frac{e}{e-1}$ -approximation with respect to the weight of a maximum weight Pareto optimal matching.
4. We finally observe that our problem has an “online” or sequential flavour similar to secretary problems⁴. Given this interpretation, we prove that our mechanism uses the best random strategy of interviewing the applicants in the sense that any other strategy would lead to an approximation ratio worse than $\frac{e}{e-1}$ (see also below under related work).

Discussion of technical contributions. Indifferences in agents’ preference lists introduce major technical difficulties. When extending SDM from strict preferences to preferences with ties, we first present an intuitive mechanism, called SDMT-1, based on the idea of augmenting paths. It is relatively easy to prove that SDMT-1 is Pareto optimal and truthful. However, it is difficult to analyse the approximation guarantee of the randomized version of SDMT-1. For this purpose we build on the primal-dual analysis of Devanur et al. [15]. They employ a linear programming (LP) relaxation of the bipartite weighted matching problem. They prove that their dual solution is feasible in expectation for the dual LP and use it to show the approximation guarantee. Towards this goal they prove two technical lemmas, a dominance lemma and monotonicity lemma. The randomized version of SDMT-1 uses random variables Y_i for each agent $i \in N$ to generate a random order in which agents are considered. Considering agent i and fixed values of the random variables Y_{-i} of all other agents, Devanur et al. [15] define a threshold y^c which as Y_i varies determines when agent i is matched (to an object) or unmatched (dominance lemma). The monotonicity lemma shows how values of the dual LP variables change when Y_i varies. To extend the definition of y^c , we need to remember the structure of all potential augmenting paths in SDMT-1, and for this purpose we introduce a second mechanism, SDMT-2. Interestingly, SDMT-2 is inspired by the idea of top trading cycle mechanisms, see, e.g., [30], however it retains the “sequential” nature of SDMT-1. The two mechanisms, SDMT-1

²A randomized mechanism is universally truthful if it is a probability distribution over truthful deterministic mechanisms. This is the strongest known notion of truthfulness for randomized mechanisms.

³A deterministic mechanism in settings with strict preferences is non-bossy if no agent can misreport his preferences in such a way that his allocation is not changed but the allocation of some other agent is changed.

⁴In the secretary problem, an administrator is willing to hire the best secretary out of n rankable applicants for a position. The applicants are interviewed one-by-one in random order. A decision about each particular applicant is to be made immediately after the interview. Once rejected, an applicant cannot be recalled. During the interview, the administrator can rank the applicant among all applicants interviewed so far, but is unaware of the quality of yet unseen applicants. The question is about the optimal strategy to maximize the probability of selecting the best applicant.

and SDMT-2, are equivalent: they match the same agents, giving them objects from the same ties. This implies that SDMT-2 is also truthful and Pareto optimal. SDMT-2 is the key to defining the threshold y^c : its running time is no worse than that of SDMT-1, but it implicitly maintains all relevant augmenting paths arising from agents' ties. We prove the monotonicity and dominance lemmas for SDMT-2 by carefully analysing the structure of frozen subgraphs that are generated from the relevant augmenting paths; here frozen roughly means that they will not change subsequently. This new approach may find further applications for primal-dual analysis of mechanisms for assignment problems where agents have indifferences. Finally, we would like to highlight that our proof of an $\frac{18}{13}$ lower bound on the approximation ratio of any universally truthful and Pareto optimal mechanism uses Yao's minmax principle and an interesting case analysis to account for all such possible mechanisms.

Related work. This work can be placed in the context of designing truthful approximate mechanisms for problems in the absence of monetary transfer [25]. Bogomolnaia and Moulin [9] designed a randomized weakly truthful and envy-free mechanism, called the Probabilistic Serial mechanism (PS), for HA with complete lists. Very recently the same authors considered the same approximation problem as ours but in the context of envy-free rather than truthful mechanisms, and for strict preference lists and unweighted agents [11]. They showed that PS has an approximation ratio of $\frac{e}{e-1}$, which is tight for any envy-free mechanism. Bhalgat et al. [8] investigated the social welfare of PS and RSDM. Tight deterministic truthful mechanisms for weighted matching markets were proposed by Dughmi and Ghosh [17] and they also presented an $O(\log n)$ -approximate random truthful mechanism for the Generalized Assignment Problem (GAP) by reducing, with logarithmic loss in the approximation, to the solution for the value-invariant GAP. Aziz et al. [5] studied notions of fairness involving the stochastic dominance relation in the context of HA, and presented various complexity results for problems involving checking whether a fair assignment exists. Chakrabarty and Swamy [13] proposed *rank approximation* as a measure of the quality of an outcome and introduced the concept of *lex-truthfulness* as a notion of truthfulness for randomized mechanisms in HA.

RSDM is related to online bipartite matching algorithms. The connection was observed by Bhalgat et al. [11], who noted the similarity between RSDM and the RANKING algorithm of Karp et al. [21]. Bhalgat et al. [11] drew a conceptual connection between a result of Karp et al. [21] — that the expected size of the matching obtained by RANKING is $\frac{e}{e-1}$ — and their own result, namely that RSDM has approximation ratio $\frac{e}{e-1}$.

The weighted version of our problem is related to two widely studied online settings, known in the literature as the online vertex-weighted bipartite matching problem [3] and secretary problems [6]. In our problem the administrator holds all the objects (they can be thought of as available positions), and all agents with unknown preference lists are applicants for these objects. Each applicant also has a private weight, which can be thought of as their quality (reflecting the fact that some of an agent's skills may not be evident from their CV, for example). However we assume that they cannot overstate their weights (skills), because they might be checked and punished. This is similar to the classical assumption of no overbidding (e.g., in sponsored search auctions). Applicants are interviewed one-by-one in a random order. When an applicant arrives he chooses his most-preferred available object and the decision as to whether it is allocated to him is made immediately, and cannot be changed in the future.

Our weighted agents correspond to weighted vertices in the vertex-weighted bipartite matching context, but our objects do not arrive online as in the setting of [3]. However, if the preference ordering of each agent in our setting, over his acceptable objects, coincides

with the arrival order of the objects in [3], then the two problems are the same. In the transversal matroid secretary problem, see, e.g., [16], objects are known in advance as in our setting, weighted agents arrive in a (uniform) random order, and the goal is to match them to previously unmatched objects. The administrator's goal is the optimal strategy, which is the random arrival order of agents, to maximize the ratio between the total weight of matched agents and the maximum weight of a matching if all the applicants' preference lists are known in advance. We show that even if the weights of all agents are the same, our algorithm uses the best possible random strategy; no other such strategy leads to better than $\frac{e}{e-1}$ -approximate matching.

Organization of the paper. The remainder of the paper is organized as follows. In Section 2 we define notation and terminology used in this paper, and show the straightforward lower bound for the approximation ratio of deterministic truthful mechanisms. SDMT-1 and SDMT-2 are presented in Sections 3 and 4 respectively, and in the latter section it is proved that the two mechanisms are essentially equivalent. The approximation ratio of $\frac{e}{e-1}$ for the randomized version of the two mechanisms is established in Section 5, whilst Section 6 contains our lower bound results. Finally, some concluding remarks are given in Section 7.

2 Definitions and preliminary observations

Let $N = \{1, 2, \dots, n_1\}$ be a set of n_1 agents and A be a set of n_2 objects. Let $n = n_1 + n_2$. Let $[i]$ denote the set $\{1, 2, \dots, i\}$. We assume that each agent $i \in N$ finds a subset of objects acceptable and has a preference ordering, not necessarily strict, over these objects. We write $a_t \succ_i a_s$ to denote that agent i *strictly prefers* object a_t to object a_s , and write $a_t \simeq_i a_s$ to denote that i is *indifferent between* a_t and a_s . We use $a_t \succeq_i a_s$ to denote that agent i either strictly prefers a_t to a_s or is indifferent between them, and say that i *weakly prefers* a_t to a_s . In some cases a weight w_i is associated with each agent i , representing the priority or importance of the agent. Weights need not be distinct. Let $W = (w_1, w_2, \dots, w_{n_1})$. To simplify definitions, we assume that all agents are assigned weight equal to 1 if we are in an unweighted setting.

We assume that the indifference relation is transitive. This implies that each agent essentially divides his acceptable objects into different bins or *indifference classes* such that he is indifferent between the objects in the same indifference class and has a strict preference ordering over these indifference classes. For each agent i , let C_k^i , $1 \leq k \leq n_2$, denote the k th indifference class, or *tie*, of agent i . We also assume that if there exists $l \in [n_2]$, where $C_l^i = \emptyset$, then $C_q^i = \emptyset$, $\forall q, l \leq q \leq n_2$. We let $L(i) = (C_1^i \succ_i C_2^i \succ_i \dots \succ_i C_{n_2}^i)$ and call $L(i)$ the preference list of agent i . We abuse notation and write $a \in L(i)$ if a appears in preference list $L(i)$, i.e., if agent i finds object a acceptable. We say that i ranks object a k th if $a \in C_k^i$. We denote by $\text{rank}(i, a)$ the rank of object a in agent i 's preference list and let $\text{rank}(i, a) = n_2 + 1$ if a is not acceptable to i . Therefore $a_t \succ_i a_s$ if and only if $\text{rank}(i, a_t) < \text{rank}(i, a_s)$, and $a_t \simeq_i a_s$ if and only if $\text{rank}(i, a_t) = \text{rank}(i, a_s)$.

Let $L = (L(1), L(2), \dots, L(n_1))$ denote the joint preference list profile of all agents. We write $L(-i)$ to denote the joint preference list profile of all agents except agent i ; i.e. $L(-i) = (L(1), \dots, L(i-1), L(i+1), \dots, L(n_1))$. Let \mathcal{L} denote the set of all possible joint preference list profiles. An instance of HA is denoted by $I = (N, A, L, W)$. We drop W and write $I = (N, A, L)$ if we are dealing with an instance where agents are not assigned weights, or equivalently if they all have the same weight. Let \mathcal{I} denote the set of all possible instances of HA.

A *matching* μ is a subset of $N \times A$ such that each agent and object appears in at most

one pair of μ . If $(i, a) \in \mu$, agent i and object a are said to be *matched together*. If $(i, a) \in \mu$ for some a , we say that i is *matched*, and *unmatched* otherwise. The definitions of *matched* and *unmatched* for an object are analogous. If agent i is matched, $\mu(i)$ denotes the object matched to i . Similarly if object a is matched, $\mu^{-1}(a)$ denotes the agent matched to a . In what follows, we will consider the following undirected graph $G = (V, E)$ where $V = (N \cup A)$ and $E = \{(i, a), i \in N, a \in L(i)\}$. We also use μ to denote a matching (in the standard graph-theoretic sense) in G . The *size* of a matching μ is equal to the number of agents matched under μ . In the presence of weights, the *weight* of a matching is equal to the sum of the weights of the matched agents.

For two given matchings μ_1, μ_2 , we will frequently use $\mu_1 \oplus \mu_2$ to denote the symmetric difference with respect to their sets of edges. An *alternating path* in G , given a matching μ_1 , is a path that consists of edges that alternately belong to μ_1 and do not belong to μ_1 . An *augmenting path* is an alternating path where the first and the last vertices on the path are unmatched in μ_1 . To *augment* along an augmenting path, given matching μ_1 , means that a new matching μ_2 is created by removing edges on the path that belong to μ_1 and adding edges on the path that do not belong to μ_1 .

A matching μ is *Pareto optimal* if there is no other matching under which some agent is better off while none is worse off. Formally, μ is *Pareto optimal* if there is no other matching μ' such that (i) $\mu'(i) \succeq_i \mu(i)$ for all $i \in N$, and (ii) $\mu'(i') \succ_{i'} \mu(i')$ for some agent $i' \in N$.

Let \mathcal{M} denote the set of all possible matchings. A *deterministic mechanism* ϕ maps an instance of HA to a matching, i.e. $\phi : \mathcal{I} \rightarrow \mathcal{M}$. Let $R : \mathcal{M} \rightarrow [0, 1]$ denote a distribution over possible matchings (which we also call a *random matching*); i.e. $\sum_{\mu \in \mathcal{M}} R(\mu) = 1$. A *randomized mechanism* ϕ is a mapping from \mathcal{I} to a distribution over possible matchings, i.e., $\phi : \mathcal{I} \rightarrow \text{Rand}(\mathcal{M})$, where $\text{Rand}(\mathcal{M})$ is the set of all random matchings. A deterministic mechanism is Pareto optimal if it returns a Pareto optimal matching. A randomized mechanism is Pareto optimal if it returns a distribution over Pareto optimal matchings.

Agents' preferences are private knowledge and an agent may prefer not to reveal his preferences truthfully if it is not in his best interests, for a given mechanism. A *deterministic mechanism* is *dominant strategy truthful* (or *truthful*) if agents always find it in their best interests to declare their preferences truthfully, no matter what other agents declare, i.e., for every agent i and every possible declared preference list $L'(i)$ for i , $\phi(L(i), L(-i)) \succeq_i \phi(L'(i), L(-i))$, $\forall L(i), \forall L(-i)$. A randomized mechanism ϕ is *universally truthful* if it is a probability distribution over deterministic truthful mechanisms.

Denote by $w(\phi(I))$ the (expected) weight of the (random) matching generated by mechanism ϕ on instance $I \in \mathcal{I}$, and by $w(I)$ the weight of a maximum weight Pareto optimal matching in I . The *approximation ratio* of ϕ is then defined as $\max_{I \in \mathcal{I}} \frac{w(I)}{w(\phi(I))}$. We give a straightforward lower bound for the approximation ratio of any deterministic truthful mechanism for HA with strict preferences.

Theorem 2.1. *No deterministic truthful Pareto optimal mechanism for HA can achieve approximation ratio better than 2. The result holds even for strict preferences.*

Proof. Consider a setting with two agents, 1 and 2, and two objects, a_1 and a_2 . Assume that both agents have weight 1 and prefer a_1 to a_2 . This setting admits two Pareto optimal matchings, both of size (weight) 2. Assume, for a contradiction, that there exists a truthful Pareto optimal mechanism ϕ with approximation ratio strictly smaller than 2. Hence, in the given setting, ϕ must pick one of the two matchings of size 2. Assume, without loss of generality, that ϕ picks $\mu = \{(1, a_2), (2, a_1)\}$. Now, assume that 1 misrepresents his acceptable objects and declares a_1 as the only object acceptable to him. As ϕ is truthful, it must not assign a_1 to 1, or else 1 finds it in his best interests to misrepresent his

preferences as he prefers his allocated object under the new setting to his allocated object under the original setting. Hence ϕ must return a matching of size at most 1 (by assigning an object to 2). However, the new setting admits a Pareto optimal matching of size 2, $\mu' = \{(1, a_1), (2, a_2)\}$. Therefore the approximation ratio of ϕ is 2, a contradiction. \square

As mentioned in Section 1, the upper bound of 2 is achievable via SDM for HA with strict preferences. If weights and ties exist, simply ordering the agents in decreasing order of their weights and running SDMT-1 (see Algorithm 1 in Section 3) or SDMT-2 (see Algorithm 2 in Section 4) gives a deterministic truthful and Pareto optimal mechanism with approximation ratio 2 (Theorem 3.6 in Section 3). This resolves the problem for deterministic mechanisms and motivates looking into relaxing our requirements. In the following sections we look for randomized truthful mechanisms that construct ‘large’ weight Pareto optimal matchings.

3 First truthful mechanism: SDMT-1

3.1 Introduction

When preferences are strict, SDM produces a Pareto optimal matching. It turns out that when indifference is allowed, finding an arbitrary Pareto optimal matching is not as straightforward as in the case of strict preferences. Take the following straightforward extension of SDM: given the agents’ preference lists, break the ties arbitrarily and then feed the resulting strict preferences to SDM. This mechanism unfortunately may fail to produce a Pareto optimal matching. To see this, consider a setting with two agents, 1 and 2, and two objects, a_1 and a_2 . Assume that 1 finds both objects acceptable and is indifferent between them, and that 2 finds only a_1 acceptable. The only Pareto optimal matching for this setting is $\mu = \{(1, a_2), (2, a_1)\}$. Assume that 1 is served first and that, after the tie-breaking stage, 1 prefers a_1 to a_2 . SDM then assigns a_1 to 1 and leaves 2 unmatched, producing a matching that is not Pareto optimal.

In Section 3.2 we introduce SDMT-1, Serial Dictatorship Mechanism with Ties, a mechanism that generalizes SDM to the case where agents’ preferences may involve ties. SDMT-1 is truthful and is guaranteed to produce a Pareto optimal matching. Further, in Section 3.3 we show that SDMT-1 is capable of generating any given Pareto optimal matching.

3.2 Mechanism SDMT-1

Let $I = (N, A, L)$ be an instance of HA, and let a fixed order σ of the agents be given. Assume, w.l.o.g., that $\sigma(i) = i$ for all agents $i \in N$. The formal description of SDMT-1 is given in Algorithm 1; an informal description follows. SDMT-1 constructs an undirected bipartite graph $G = (V, E)$ where $V = N \cup A$ and the set of edges E changes during the execution of SDMT-1. The mechanism proceeds in n_1 phases and returns a matching μ . Each phase of the algorithm corresponds to one iteration of the for loop in Algorithm 1. Let μ_i denote the matching at the end of phase i . Initially $E = \emptyset$ and $\mu = \emptyset$. In each phase i , agent i is considered and the objects in agent i ’s preference list are examined in the order of the indifference classes they belong to. When objects $a \in C_\ell^i$ are examined, edges (i, a) are provisionally added to E for all $a \in C_\ell^i$. We then check whether there is an augmenting path in G that starts from agent i . If such an augmenting path exists, we augment along that path and modify μ accordingly. This would mean that agent i is assigned a and every other agent j matched under μ_{i-1} is assigned an object that he ranks

Algorithm 1: Serial Dictatorship Mechanism with Ties, version 1 (SDMT-1)

Input: Agents N ; Objects A ; Preference list profile L ; An order of agents σ
Output: Matching μ
Let $G = (N \cup A, E)$, $E \leftarrow \emptyset$, $\mu \leftarrow \emptyset$.
for each agent $i \in N$ in the order of σ **do**
 Let $\ell \leftarrow 1$
 Step (*): **if** $C_\ell^i \neq \emptyset$ **then**
 $E \leftarrow E \cup \{(i, a) : a \in C_\ell^i\}$; // all new edges are non-matching edges
 if there is an augmenting path from i in G **then**
 augment along this path and update μ accordingly; // i is provisionally
 allocated some $a \in C_\ell^i$ and (i, a) is now a matching edge
 end
 else
 $E \leftarrow E \setminus \{(i, a) : a \in C_\ell^i\}$
 $\ell \leftarrow \ell + 1$; Go to Step (*)
 end
 end
end
Return μ ; //each matched agent is allocated his matched object

in the same indifference class as $\mu_{i-1}(j)$. Otherwise—if G admits no augmenting path, edges (i, a) are removed from E for all $a \in C_\ell^i$.

Notice that, at any stage of the mechanism, an edge (i, a) belongs to E if and only if either agent i is matched in μ and $a \simeq_i \mu(i)$ or SDMT-1 is at phase i and examining the indifference class to which a belongs. Therefore, it is fairly straightforward to observe the following.

Observation 3.1. *At the end of phase i of SDMT-1, if agent i is assigned no object then he will be assigned no object when SDMT-1 terminates. Otherwise, if i is provisionally assigned an object a , then he will be allocated an object that he ranks the same as a in the final matching.*

Before proceeding to prove our main claim, that SDMT-1 is truthful and produces a Pareto optimal matching, let us discuss a relevant concept that is both interesting in its own right and useful in the proofs that follows. In practice agents may have priorities and the mechanism designer may wish to ensure that the agents with higher priorities are served before satisfying those with lower priorities. Roth et al. [28] studied this concept under the term *priority matchings* in the case where each agent's preference list is one big tie. This work was motivated by the kidney exchange problem in which patients are assigned priorities based on various criteria; e.g. children and hard-to-match patients have higher priorities. Prior to Roth et al. [28], Svensson [31] studied a similar concept under the name *queue allocation* in a setting similar to ours. We formally define this concept using the terminology *strong priority matching*, reflecting both the definition in [28] and the fact that preference lists are more general than single ties.

Fix an ordering of the agents $\sigma = i_1, \dots, i_{n_1}$. For each matching μ , the *signature* of μ , denoted by $\rho(\mu)$, is a vector $\langle \rho_1, \dots, \rho_{n_1} \rangle$ where for each $k \in [n_1]$, $\rho_k = \text{rank}(i_k, \mu(i_k))$ if i_k is matched under μ , and $\rho_k = n_2 + 1$ otherwise. A matching μ is a *strong priority matching (SPM)* w.r.t. σ if μ has the lexicographically smallest signature. It is easy to see that a given setting may admit more than one SPM w.r.t. σ , but all of them have the same signature. When σ is fixed and known, we simply say that μ is a strong priority matching.

Theorem 3.2. *The matching produced by SDMT-1 is a strong priority matching w.r.t. σ .*

Proof. Assume, for a contradiction, that the claim does not hold. Hence μ is not a SPM in I . Let μ^* be a SPM in I . Let z be the lowest indexed agent (and therefore first in σ) who prefers his partner under μ^* to his partner under μ , i.e. $\mu^*(z) \succ_z \mu(z)$ and $\mu^*(x) \simeq_x \mu(x)$, $\forall x < z$ (we denote this deduction by **D1**). Therefore, in phase z of SDMT-1 no augmenting path has been found starting from (z, a) where $a \succeq_z \mu^*(z)$ (we denote this deduction by **D2**). Also, it follows from **D1** and Observation 3.1 that, $\mu^*(x) \simeq_x \mu_{z-1}(x)$, $\forall x < z$ (we denote this deduction by **D3**). Let G^* denote the graph G in phase z during the examination of the indifference class to which $\mu^*(z)$ belongs. By **D2**, G^* must admit no augmenting path w.r.t. μ_{z-1} . We show, however, that G^* admits an augmenting path starting from z . To see this note that, by **D1** and the construction of edges E , edges $(x, \mu^*(x))$ belong to G^* . If $\mu^*(z)$ is unmatched in μ_{z-1} then $(z, \mu^*(z))$ constitutes an augmenting path of size 1 in G^* . Otherwise, let i_1 denote the partner of $\mu^*(z)$ under μ_{z-1} (note that $i_1 < z$). It follows, from **D3** and the construction of E , that i_1 is matched under μ^* . If $\mu^*(i_1)$ is unmatched under μ_{z-1} then we have found an augmenting path of length 3. Otherwise, we continue with this argument. As there are a finite number of agents and objects, we have to reach an object a who is unmatched under μ_{z-1} , hence exposing an augmenting path in G^* , a contradiction. \square

Corollary 3.3. *The matching produced by SDMT-1 is a Pareto optimal matching.*

Proof. By Theorem 3.2, SDMT-1 produces a strong priority matching. It follows from Theorems 1 and 2 in [31] that any strong priority matching is a Pareto optimal matching. Hence, the matching produced by SDMT-1 is a Pareto optimal matching. \square

SDMT-1 is truthful, no matter which augmenting path is selected in each phase of the mechanism, as the next result shows.

Theorem 3.4. *The mechanism SDMT-1 is truthful.*

Proof. Assume, for a contradiction, that the claim does not hold. Assume that agent i benefits from misrepresenting his preferences and reporting $L'(i)$ instead of $L(i)$. Let $L' = (L'(i), L(-i))$. Let μ denote the matching returned by SMDT-1 on instance $I = (N, A, L)$ —i.e. the instance in which agent i reports truthfully, and μ' denote the matching returned on instance $I' = (N, A, L')$. Let μ_ℓ and μ'_ℓ , $\forall \ell \in N$, denote the matchings generated at the end of phase ℓ given I and I' respectively. Let $a = \mu(i)$ and $a' = \mu'(i)$. Let $u = \text{rank}(i, a)$ and $u' = \text{rank}(i, a')$. It follows that $a' \succ_i a$ and $u' < u$. By Observation 3.1 and the construction of L' , it must be the case that $\mu_\ell(z) \simeq_z \mu'_\ell(z)$ $\forall \ell, z$, $1 \leq z \leq \ell < i$. Furthermore, $\mu_i(i) \simeq_i a$ and $\mu'_i(i) \simeq_i a'$. Consider the execution of SMDT-1 on I . Let G^* denote the graph G in phase i during the examination of $C_{u'}^i$. As $u' < \text{rank}(i, \mu(i))$, G^* must not admit an augmenting path w.r.t. μ_{i-1} . We show, however, that G^* admits an augmenting path. If a' is unmatched in μ_{i-1} then (i, a') constitutes an augmenting path of size 1 in G^* . Otherwise, let i_1 denote the partner of a' under μ_{i-1} (note that $i_1 < i$). The rest of the argument is similar to that of presented in the proof of Theorem 3.2, leading to the conclusion that G^* admits an augmenting path, a contradiction. \square

We now show a bound on the time complexity of SDMT-1. Let γ denote the size of the largest indifference class for a given instance I .

Theorem 3.5. *SDMT-1 terminates in time $O(n_1^2 \gamma)$.*

Proof. For each agent i matched under μ , let ℓ_i denote the length of the indifference class to which $\mu(i)$ belongs. Let $|L(i)|$ denote the length of agent i 's preference list, $\forall i \in N$. Search for an augmenting path in a graph $G = (V, E)$ can be done in time $O(|E|)$ using Depth First Search (DFS). Hence the search for an augmenting path in each phase i can be done in time $O(\ell_1 + \ell_2 + \dots + \ell_{i-1} + |L(i)|)$. Therefore SDMT-1 terminates in time $O(n_1 \cdot \ell_1 + (n_1 - 1)\ell_2 + \dots + \ell_{n_1} + \sum_{i \in N} |L(i)|)$. However, $\ell_i \leq \gamma$, $\forall i \in N$, therefore $n_1 \cdot \ell_1 + (n_1 - 1)\ell_2 + \dots + \ell_{n_1} + \sum_{i \in N} |L(i)| \leq n_1^2 \gamma + m$, where m is the number of (agent,object) acceptable pairs. Hence SDMT-1 terminates in time $O(n_1^2 \gamma)$. \square

As noted earlier, any Pareto optimal matching is at least half the size of a maximum size such matching (see, e.g., [2]). Hence SDMT-1 obviously achieves approximation ratio 2 when we are concerned with the cardinality of the matching. We next show that, when agents are assigned arbitrary weights, SDMT-1 achieves the same approximation ratio (relative to a maximum weight Pareto optimal matching) if the agents are ordered in σ in non-increasing order of their weights.

Theorem 3.6. *SDMT-1 achieves approximation ratio of 2 relative to the size of a maximum weight Pareto optimal matching, if the agents are ordered in σ in non-increasing order of their weights.*

Proof. Given an instance I , let μ be the matching produced by SDMT-1 and let OPT be a maximum weight Pareto optimal matching in I . List the agents matched under each of these matchings in non-increasing order of weight. Let i_1, \dots, i_k denote such an order under μ , and let i'_1, \dots, i'_l denote such an order under OPT .

Take any agent i'_x who is matched under OPT , to say a_x , but not matched under μ (if no such agent exists then μ is itself a maximum weight Pareto optimal matching). Note that, as μ is Pareto optimal, a_x must be matched under μ , for otherwise $\mu \cup \{(i'_x, a_x)\}$ Pareto dominates μ . As SDMT-1 generates a strong priority matching w.r.t. σ (Theorem 3.2) and agents are listed in non-increasing order of weight under σ , it follows that a_x must be allocated in μ to an agent i_y who has at least as large a weight as i'_x (for otherwise $(\mu \setminus \{(i_y, a_x)\}) \cup \{(i'_x, a_x)\}$ has a lexicographically smaller signature than μ , a contradiction).

We claim that i_y must be matched under OPT as well, as otherwise $OPT \setminus \{(i'_x, a_x)\} + \{(i_y, a_x)\}$ has a higher weight than OPT . (Note that a maximum weight Pareto optimal matching must be a maximum weight matching as well.) Hence we have established that, for each agent i'_x matched under OPT but not matched under μ , there exists a unique agent i_y , with weight at least as large as that of i'_x , who is matched under μ . Thus if N_1 is the set of agents matched in OPT and N_2 is the set of agents matched in μ , it follows that $wt(N_2) \geq wt(N_1 \setminus N_2)$, where $wt(N')$ is the sum of the weights of the agents in N' , for $N' \subseteq N$. Also $wt(N_2) = wt(N_2 \setminus N_1) + wt(N_2 \cap N_1) \geq wt(N_2 \cap N_1) = wt(N_1) - wt(N_1 \setminus N_2) \geq wt(N_1) - wt(N_2)$, hence the result. \square

3.3 Enumerating all Pareto optimal matchings

It is known that (see, e.g., [2]), in the case of strict preferences, not only we can find a Pareto optimal matching using SDM, but we can also generate *all* Pareto optimal matchings by executing SDM on all possible permutations of the agents. In other words, given any Pareto optimal matching μ , there exists an order of the agents such that executing SDM on that order returns μ . In this section we seek a similar characterization of Pareto optimal matchings in the case of preferences with ties.

Theorem 3.7. *Any Pareto optimal matching is a strong priority matching for some ordering σ of the agents.*

Proof. This proof makes use of the characterization of Pareto optimal matchings in instances of HA (potentially with ties) given in [22] in terms of a number of graph-theoretic structures. In particular, a given matching μ is Pareto optimal if and only if μ admits no *alternating path coalition*, no *augmenting path coalition*, and no *cyclic coalition* (see [22]); we will now define these terms.

An *alternating path coalition* w.r.t. μ compromises a sequence $P = \langle i_0, i_1, \dots, i_{r-1}, a_k \rangle$, for some $r \geq 1$, where i_j is a matched agent ($0 \leq j \leq r-1$) and a_k is an unmatched object. If $r = 1$ then i_0 prefers a_k to $\mu(i_0)$. Otherwise, if $r \geq 2$, i_0 prefers $\mu(i_1)$ to $\mu(i_0)$, i_j weakly prefers $\mu(i_{j+1})$ to $\mu(i_j)$ ($1 \leq j \leq r-2$), and i_{r-1} weakly prefers a_k to $\mu(i_{r-1})$.

An *augmenting path coalition* w.r.t. μ compromises a sequence $P = \langle i_0, i_1, \dots, i_{r-1}, a_k \rangle$, for some $r \geq 1$, where i_0 is an unmatched agent and a_k is an unmatched object. If $r = 1$ then i_0 finds a_k acceptable. Otherwise, if $r \geq 2$, i_j is a matched agent ($1 \leq j \leq r-1$), i_0 finds $\mu(i_1)$ acceptable, i_j weakly prefers $\mu(i_{j+1})$ to $\mu(i_j)$ ($1 \leq j \leq r-2$), and i_{r-1} weakly prefers a_k to $\mu(i_{r-1})$.

A *cyclic coalition* w.r.t. μ compromises a sequence of applicants $P = \langle i_0, i_1, \dots, i_{r-1} \rangle$, for some $r \geq 2$, all matched in μ , such that i_j weakly prefers $\mu(i_{j+1})$ to $\mu(i_j)$ for each j ($0 \leq j \leq r-1$), and i_j prefers $\mu(i_{j+1})$ to $\mu(i_j)$ for some j ($0 \leq j \leq r-1$) (all subscripts are taken modulo r when reasoning about cyclic coalitions).

Let a Pareto optimal matching μ be given. Let $G = (V, E)$ be the *envy graph* for μ , defined as follows. In the graph, $V = N$ and there is a directed edge from agent i to agent k iff i weakly prefers $\mu(k)$ to $\mu(i)$. Every edge is colored. An edge (i, k) is colored green if $\mu(k) \simeq_i \mu(i)$, and is colored red otherwise—i.e. if $\mu(k) \succ_i \mu(i)$. We claim that all the edges in every strongly connected component (SCC) of G are green (we denote this claim by **C1**). To see this, note that by the definition of strongly connected components, there is a path from every node in a given component to every other node in the component. Hence, if there is a red edge (i, k) in a SCC, then there must be a cycle with a red edge in the SCC (as there must be a path from k to i). A cycle with at least one red edge corresponds to a cyclic coalition and hence μ could have not been a Pareto optimal matching, a contradiction.

Create graph $G' = (V', E')$ as follows. There is a vertex in V' for each SCC of G , and there is a directed edge in G' from v_r to v_s , $v_r, v_s \in V'$ and $v_r \neq v_s$, iff there is an edge in G from i to k for some i and k that belong to the SCCs of v_r and v_s . It follows from the definition of strongly connected components that G' is a DAG. Hence G' admits a topological ordering. Let X be a reversed topological ordering of G' . Let $\sigma = i_1, \dots, i_{n_1}$ be an ordering of all the agents that is consistent with X . That is, for every two agents i_j and i_r , $1 \leq j < r \leq n_1$, the corresponding SCC of i_j appears in X no later than the corresponding SCC of i_r . (The order of the agents belonging to the same SCC can be determined arbitrarily.) We prove that μ is a strong priority matching w.r.t. σ .

Assume, for a contradiction, that our claim does not hold. That is, μ is not a strong priority matching w.r.t. σ . Hence there must exist another matching μ' which has a lexicographically smaller signature than μ ; i.e. $\rho(\mu') < \rho(\mu)$ (we denote this deduction by **A1**). Let i_j be the highest priority agent, w.r.t. σ , such that i_j prefers his partner under μ' to his partner under μ ; i.e. $\mu'(i_j) \succ_{i_j} \mu(i_j)$ (we denote this assumption by **A2**). Note that $\mu'(i_j)$ must be matched in μ or else μ admits an alternating coalition, namely $P = \langle i_j, \mu'(i_j) \rangle$ —as i_j prefers unmatched object $\mu'(i_j)$ to its partner $\mu(i_j)$ —and hence not Pareto optimal. So $\mu'(i_j)$ is matched under μ to, say, i_k . Following **A2**, there must be a red edge from i_j to i_k in the envy graph G . Therefore, i_k must have a higher priority than i_j according to σ (note that, by **C1**, i_j and i_k cannot belong to the same SCC). It then follows from **A1** and **A2** that i_k is matched under μ' and ranks its partners under μ and μ' the same; i.e. $\mu(i_k) \simeq_{i_k} \mu'(i_k)$. Now, $\mu'(i_k)$ must be matched in μ or else

there is an alternating path coalition in μ , namely $P = \langle i_j, i_k, \mu'(i_k) \rangle$, and hence μ is not Pareto optimal. Also, i_k cannot be matched to $\mu(i_j)$ or else there is a cyclic coalition in μ , namely $P = \langle i_j, i_k \rangle$, and hence μ is not Pareto optimal. So $\mu'(i_k)$ is matched under μ to, say, i_r , $i_r \neq i_j$. Also, there is a green edge from i_k to i_r in G , since $\mu'(i_k) = \mu(i_r)$ and $\mu(i_k) \simeq_{i_k} \mu'(i_k)$. We claim that i_r must have a higher priority than i_j (we denote this claim by **C2**). To see this, first of all note that there cannot exist a path between i_r and i_j in G or else G admits a cycle with at least one red edge, namely (i_j, i_k) , and hence μ admits a cyclic coalition and is not Pareto optimal. Now, to complete the proof of **C2**, we consider two cases regarding whether there is a path from i_r to i_k or not.

- Case 1: *There is a path from i_r to i_k in G .* Since there is an path from i_k to i_r , then i_r and i_k must belong to the same strongly connected component. Therefore since i_k has a higher priority than i_j under σ , so must i_r .
- Case 2: *There is no path from i_r to i_k in G .* Therefore i_r and i_k belong to two different SCCs. Since there is an path from i_k to i_r in G , there is an edge in G' from the SCC corresponding to i_k to the SCC corresponding to i_r . Therefore i_r must have a higher priority than i_k under X and thus under σ as well. Hence, by transitivity, i_r has a higher priority than i_j under σ .

So far we have established that i_r has a higher priority than i_j . Now, using a similar argument as for i_k , we can show that i_r must be matched under μ' and must rank his partners under μ and μ' the same. Also, again using a similar argument as for i_k , $\mu'(i_r)$ is not the same object as $\mu(i_j)$ and $\mu'(i_r)$ must be matched in μ or else there is an alternating path coalition in μ contradicting the Pareto optimality of μ . So $\mu'(i_r)$ is matched under μ , to say i_q . Using exactly the same argument as we used for i_r we can show that i_q also has a higher priority than i_j . We can keep repeating this argument and every time we have to reach a new agent with a higher priority than i_j . However, there are a bounded number of agents and hence a bounded number of agents with higher priority than i_j , a contradiction. \square

Following Theorem 3.7, enumerating all Pareto optimal matchings is equivalent to enumerating all matchings μ such that μ is a strong priority matching w.r.t. some ordering of the agents. Let μ^* be a matching returned by SDMT-1 for a given priority ordering of the agents $\sigma = i_1, \dots, i_{n_1}$. Recall that all strong priority matchings w.r.t. σ have the same signature as $\rho(\mu^*)$. It hence follows that,

Proposition 3.8. *A matching μ is a SPM w.r.t. σ if and only if*

- *the same set of agents are matched under both μ and μ^* , and*
- *each matched agent i is matched under μ to an object that he ranks the same as $\mu^*(i)$; i.e., $\text{rank}(i, \mu(i)) = \text{rank}(i, \mu^*(i))$.*

Let $G(\sigma) = (V, E)$ be a graph where $V = N^* \cup A$ where N^* is the set of agents that are matched under μ^* . There is an edge between an agent $i \in N^*$ and an object $a \in A$ if and only if i ranks a the same as $\mu(i)$. It is then easy to see that,

Proposition 3.9. *A matching μ is a strong priority matching w.r.t. σ if and only if it is a maximum cardinality matching in $G(\sigma)$.*

Thus to enumerate all strong priority matchings w.r.t. σ we need only to enumerate all maximum cardinality matchings w.r.t. $G(\sigma)$. The latter can be achieved in $O(|V|)$ time per matching (see [32]).

Finally, we show that SDMT-1 is capable of producing any given Pareto optimal matching.

Theorem 3.10. *Any Pareto optimal matching can be generated by some execution of SDMT-1.*

Proof. Let μ be a Pareto optimal matching for an instance I of HA. By Theorem 3.7 μ is a strong priority matching for some ordering σ of the agents. Execute SDMT-1 given σ as follows. At each phase i , choose $(i, \mu(i))$ as the augmenting path. Notice that since μ is a matching, both i and $\mu(i)$ must be unmatched at the beginning of phase i . Furthermore, since μ is a SPM w.r.t σ , there cannot be an augmenting path from i to an object that i prefers to $\mu(i)$. \square

4 Randomized mechanism with weights and ties

4.1 Online interpretation of our problem

In this section we will analyse our mechanism for the weighted version of our problem. We can interpret it as having an “online” flavour in the following way. An administrator holds all the objects, and all agents with unknown preference lists are applicants for these objects. We assume that weights are private information of each agent, but that they cannot overstate their weights. Applicants are interviewed one-by-one in a random order. A decision about each particular applicant is to be made immediately after the interview. During the interview, the applicant selects his favorite (because we only consider truthful mechanisms) object among the available remaining objects if there exists one in his preference list and is allocated (matched to) that object.⁵ This applicant will not be interviewed again. The administrator can know the number of matched applicants interviewed so far, but is unaware whether yet unseen applicants will be matched or not. Our goal is to find the optimal strategy, which is the random arrival order of agents, that maximizes the ratio between the total weight of matched agents and the maximum weight Pareto optimal matching if all the applicants’ preference lists are known in advance.

Our algorithm in the next section is truthful with respect to agents’ preferences and weights (under no-overbidding assumption) and provides an $\frac{e}{e-1}$ -approximate Pareto optimal matching. We will show in Section 6 that, even if the weights of all agents are the same our algorithm uses the best possible random strategy – no other such strategy leads to better than $\frac{e}{e-1}$ -approximate matching.

4.2 Second truthful mechanism: SDMT-2

The approximation ratio analysis of the randomized version of SDMT-1 is complex, because it requires additional information which is not maintained by SDMT-1. For the sake of the analysis, we introduce a variant of SDMT-1, called SDMT-2. After introducing some terminology we present SDMT-2, and then establish the equivalence between SDMT-1 and SDMT-2. Pareto optimality and truthfulness of SDMT-2 will then follow from this equivalence and these same two properties of SDMT-1. We will prove that the randomized version of SDMT-2 is $\frac{e}{e-1}$ -approximate. By the equivalence of the two algorithms, a randomized version of SDMT-1 has the same approximation ratio.

Let $a_1 \succ a_2 \succ \dots \succ a_{n_2}$ be a common order of all the objects; it will be used to break possible ties in SDMT-2). In the course of our algorithm agents will be (temporarily)

⁵We can extend this setting to the case where the administrator can decide whether to let the applicant select his favorite object or to reject this applicant, meaning that the applicant gets nothing. In this more general problem, it is not difficult to prove that for any fixed order of the applicants, the decision that the administrator does not reject any applicant will maximize the number of matched applicants. Therefore, this more general problem is reduced to the setting where the administrator lets each applicant select his favorite object, and therefore our lower bound from Section 6 also applies to this setting.

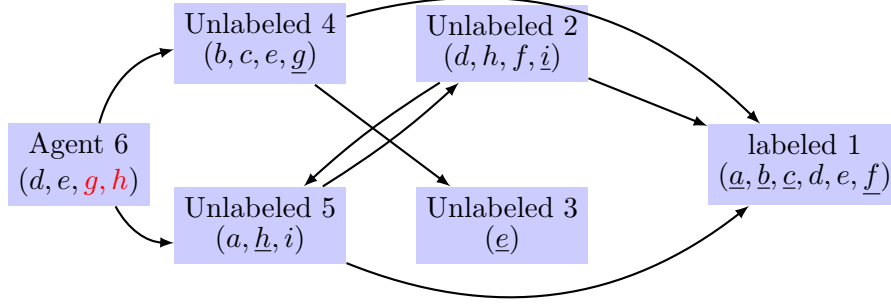


Figure 1: The trading graph $TG(6, \{g, h\}, [5])$, \underline{h} denotes h is owned currently by the agent. The common order is $a \succ b \succ c \succ d \succ e \succ f \succ g \succ h \succ i$.

allocated subsets of objects from their preference list. When an agent is allocated a subset of objects we say that he *owns* these objects. Let $S \subseteq N$ and suppose that some of the agents in S have been allocated some objects and the allocated objects to each agent appear in the same indifference class of this agent. At any time during the execution of the algorithm, each agent who is allocated more than one object is called *labeled* and *unlabeled* otherwise. Let $i \in N$, and let $B \subseteq L(i)$ be such that i is not allocated any object in B . The *trading graph* (TG) is a directed graph $TG(i, B, S)$ with $\{i\} \cup S$ as the set of nodes, and arcs defined as follows: Let agent i point to any agent in S who owns any object in B . For each unlabeled agent, e.g., $j \in S$, to which i points, suppose the current object allocated to j is in j 's k th indifference class C_k^j . Then let j point to any agent in S who currently owns any object in C_k^j not owned by j . Continue this process for the new pointed-to and unlabeled agents until no agent in S needs to point to the other agents. See Figure 1 for how $TG(6, \{g, h\}, [5])$ is constructed: agent 6 points to agent 4 and 5 since currently agent 4 owns g and agent 5 owns h ; then, as agent 4 is unlabeled, agent 4 points to agents 3 and 1 since agent 3 owns e and agent 1 owns b and c ; similarly, agent 5 points to agents 1 and 2; agent 1 is labeled so only agent 2 points to agents 1 and 5.

Let $H = \{a \in L(i) \mid \text{there is a (directed) path from } i \text{ to a labeled agent in } TG(i, a, S)\}$. Note that, as labeled agents do not point to any agents, no intermediate agent on a directed path is labeled. Note that H may be empty, and it can be found by breadth first search (BFS). If $H \neq \emptyset$, let ℓ be the highest indifference class of i with $H \cap C_\ell^i \neq \emptyset$. Define $\max TG(i, L(i), S)$ to be the highest order object in $H \cap C_\ell^i$ (e.g., in Fig. 1, $\max TG(6, \{d, e, g, h\}, [5]) = \{g\}$). If $\max TG(i, L(i), S) \neq \emptyset$, then there is a path from i to a labeled agent in $TG(i, a, S)$, which can be found by BFS. Suppose the path is $(i_0, i_1, i_2, \dots, i_k)$, where $i_0 = i$ and only i_k is labeled. Now denote $Trading(i, a, S)$ to be a procedure that allocates the object owned by i_{s+1} to i_s , for $s = 0, 1, \dots, k-1$. Note that i_k may own more than one object for which i_{k-1} has pointed to i_k . In this case, the highest order object among such objects is allocated to i_{k-1} . After trading, if i_k still owns more than one object, keep i_k labeled and unlabeled i_k otherwise. In Fig. 1, $Trading(6, g, [5])$ allocates g to agent 6 and b to agent 4, since $b \succ c$, and keeps agent 1 labeled. Note that $C_{n_2+1}^i = \emptyset, \forall i \in [n_1]$. With these preliminaries, we present our algorithm SDMT-2 (Algorithm 2). In the following, we will refer to k th iteration of the “for loop” in SDMT-2 as k th loop. Observe that in the k th loop, j_1 is the highest indifference class of i where i can obtain unallocated objects, and j_2 is the highest indifference class of i where i can

Algorithm 2: Serial Dictatorship Mechanism with Ties, version 2 (SDMT-2)

Input: Agents N ; Objects A ; Preference list profile L ; An order of agents σ ,
w.l.o.g. let $\sigma(i) = i, \forall i \in N$
Output: Matching
Let $A_1 \leftarrow A$.
for each agent $i \in N$ in the order of σ **do**
 Define $j_1 = \begin{cases} \min\{j : A_i \cap C_j^i \neq \emptyset\} & \text{if } A_i \cap L(i) \neq \emptyset; \\ n_2 + 1 & \text{otherwise.} \end{cases}$
 $j_2 = \begin{cases} \min\{j : \max TG(i, L(i), [i-1]) \in C_j^i\} & \text{if } \max TG(i, L(i), [i-1]) \neq \emptyset; \\ n_2 + 1 & \text{otherwise.} \end{cases}$
 if $j_1 \leq j_2$ **then**
 Allocate all the objects in $A_i \cap C_{j_1}^i$ to i ; Label i if $|A_i \cap C_{j_1}^i| \geq 2$;
 $A_{i+1} \leftarrow A_i \setminus (A_i \cap C_{j_1}^i)$
 end
 else
 $Trading(i, \max TG(i, L(i), [i-1]), [i-1])$; $A_{i+1} \leftarrow A_i$
 end
end
For each labeled agent, allocate to him the highest order object he currently owns.
For each unlabeled agent, if he currently owns an object, allocate it to him.
Output the matching.

obtain objects from the allocated objects without hurting the agents prior to i .

Observation 4.1. For each agent i , after i 's turn in "for loop" of SDMT-2, if i is allocated no object, then he will be allocated no object when SDMT-2 terminates. Otherwise, if i is provisionally allocated some objects in his turn, then in the final matching he will be allocated an object in the same indifference class as his initially allocated objects.

Observation 4.2. For each agent i , after i 's turn, if i is allocated an object $a \in C_j^i$, then all the objects in $\cup_{k=1}^j C_k^i$ have been allocated to either i or to some agents prior to i . Once an object is allocated, it remains allocated until the end of the for loop.

Now we establish the equivalence of SDMT-1 and SDMT-2.

Theorem 4.3. Given the same input, SDMT-1 and SDMT-2 match the same set of agents. Furthermore, for each matched agent i , the object allocated to i in SDMT-1 is in the same indifference class of i as the object allocated to him in SDMT-2.

Proof. We will prove the following two facts inductively which simply implies the conclusion of Theorem 4.3. W.l.o.g. suppose the order of agents is $\sigma(i) = i, \forall i \in N$. Until the step i ,

1. for each agent $k \leq i-1$, the allocated objects of SDMT-1 and SDMT-2 to k are in the same indifference class, (if one of them is empty, the other is empty as well)
2. for each $\ell \leq n_2$, and $a \in C_\ell^i$, there is an augmenting path starting from (i, a) in SDMT-1 if and only if a is unallocated in SDMT-2 or there is a path from i to a labeled agent in $TG(i, a, [i-1])$ in SDMT-2.

Consider the base case, for agent 1, obviously property 1 is true since they are all empty sets. if $L(1) \neq \emptyset$, then there is an augmenting path from (i, a) in SDMT-1 and a is

unallocated.

Suppose property 1 and 2 is true for all the steps $k \leq i - 1$, we now prove that it is true for step i . For property 1, by inductive hypothesis, property 1 holds for any $k \leq i - 2$, since property 2 holds for agent $i - 1$ by inductive hypothesis, the objects allocated to agent $i - 1$ in SDMT-1 and SDMT-2 will be in the same indifference class, thus, property 1 holds for step i . Now property 2 will be proved true for agent i , for each $\ell \leq n_2$, and $a \in C_\ell^i$.

For \Rightarrow direction, if there is an augmenting path starting from (i, a) in SDMT-1, and if a is allocated previously in SDMT-2, suppose the new matching generated in SDMT-1 due to the augmenting path is $(k, \mu(k))$, $k \leq i$, where $\mu(i) = a$. By property 1 of inductive hypothesis and Observation 4.2, all the objects in $\{\mu(k), k \leq i\}$ have been allocated to some agents $k \leq i - 1$ in SDMT-2. For object b , we use $\nu^{-1}(b)$ to denote the agent whom b is allocated to in SDMT-2. Now consider the following path in $TG(i, a, [i - 1])$: let $i_1 = \nu^{-1}(\mu(i))$, if i_1 is labeled then we are done, otherwise, let $i_2 = \nu^{-1}(\mu(i_1))$. If i_2 is labeled, then we are done, otherwise continue this process. Finally, we either reach a labeled agent or the path has passed through all the agents in $[i - 1]$. In the latter case, note that there are i objects ($\{\mu(k), k \leq i\}$) allocated to $i - 1$ agents ($[i - 1]$) in SDMT-2 and the last agent must therefore be allocated to at least two objects, which becomes the former case.

For \Leftarrow direction, now suppose a is unallocated or there is a path from i to a labeled agent in $TG(i, a, [i - 1])$ in SDMT-2. Suppose a is allocated and there is a path from i to a labeled agent in $TG(i, a, [i - 1])$. Then by *Trading*($i, a, [i - 1]$), we can make all the agents $k \leq i$ allocated at least one object and i is allocated a . Let us select any matching, e.g., $M = \{(k, \nu(k)), k \leq i\}$, where $\nu(i) = a$ (we can also select such a matching if a is unallocated in SDMT-2). Suppose the matching generated after step $i - 1$ in SDMT-1 is $M' = \{(k, \mu(k)), k \leq i - 1\}$. By property 1 of inductive hypothesis, we know $\mu(k)$ and $\nu(k)$ are in the same indifference class of agent k , for any $k \in [i - 1]$. Now consider $M \oplus M'$, which consists of alternating paths and cycles. Then a connected component of $M \oplus M'$ that contains $(i, \nu(i))$ must be an odd length alternating path in $M \oplus M'$ w.r.t. M' , implying an augmenting path starting from (i, a) in SDMT-1. The argument showing that the connected component that contains $(i, \nu(i))$ must be an odd length alternating path is as follows. If $\nu(i) = a$ is unallocated in SDMT-1, then $(i, \nu(i))$ is an odd length alternating path, otherwise suppose $i_1 = \mu^{-1}(\nu(i))$, then consider whether $\nu(i_1)$ is allocated or not in SDMT-1. If not we get an odd path $(i, \nu(i), i_1, \nu(i_1))$. Otherwise continue the search, and let $i_2 = \mu^{-1}(\nu(i_1))$, then consider whether $\nu(i_2)$ is allocated or not in SDMT-1. If not we get an odd length path $(i, \nu(i), i_1, \nu(i_1), i_2, \nu(i_2))$, and so on. Finally, we will get an odd length alternating path starting from $(i, \nu(i)) = (i, a)$ w.r.t. M' , which is indeed an augmenting path starting from (i, a) in SDMT-1. This concludes the induction. \square

It is easy to see that both SDMT-1 and SDMT-2 reduce to SDM if all agents have strict preference over objects. We say an agent (respectively, an object) is *frozen* if the allocation of this agent (respectively, object) remains the same until the end of SDMT-2.

Theorem 4.4. *SDMT-2 is truthful, Pareto optimal, and terminates in $O(n_1^2 \gamma)$ running time.*

Proof. The first two properties follow from the equivalence between SDMT-1 and SDMT-2 (Theorem 4.3) and the Pareto optimality (Corollary 3.3) and truthfulness (Theorem 3.4) of SDMT-1. We provide a direct proof here.

Denote ϕ as the mechanism of SDMT-2, and $\phi_i(L(i), L(-i))$ denotes the allocation of agent i under preference lists $(L(i), L(-i))$. First, we show ϕ is truthful. For any agent i , suppose i 's preference list is $L(i)$. Given preference lists L , suppose i changes his preference

list to $L'(i)$. If $\phi_i(L(i), L(-i)) = \emptyset$, then for any $a \in L(i)$, a must have been allocated to an agent $k \leq i-1$ (otherwise, $A_i \cap L(i) \neq \emptyset$, i will be allocated to at least one object by SDMT-2, a contradiction). $\max TG(i, L(i), [i-1])$ must also be empty set by SDMT-2. In other words, the objects of $L(i)$ are frozen at the beginning of i th loop. Now consider SDMT-2 running on $(L'(i), L(-i))$ and note Observation 4.1, the objects of $L(i)$ are frozen at the beginning of i th loop. Thus, either $\phi_i(L'(i), L(-i))$ is empty or an object not in $L(i)$, in either case, $\phi_i(L(i), L(-i)) \succeq_i \phi_i(L'(i), L(-i))$. Now suppose $\phi_i(L(i), L(-i)) \neq \emptyset$ and let $a = \phi_i(L(i), L(-i)) \in C_j^i$. Obviously, if $j = 1$, we have $\phi_i(L(i), L(-i)) \succeq_i \phi_i(L'(i), L(-i))$. Now suppose $j > 1$, then by SDMT-2, we know that $A_i \cap (\cup_{k=1}^{j-1} C_k^i) = \emptyset$ and for any $a \in \cup_{k=1}^{j-1} C_k^i$, there is no path from a to a labeled agent in $TG(i, a, [i-1])$ (otherwise, agent i will be allocated to an object in $\cup_{k=1}^{j-1} C_k^i$). The objects in $\cup_{k=1}^{j-1} C_k^i$ have been frozen at the beginning of i th loop. Hence, by SDMT-2 and Observation 4.1, $\phi_i(L'(i), L(-i))$ will not be an object in $\cup_{k=1}^{j-1} C_k^i$. Therefore, $\phi_i(L(i), L(-i)) \succeq_i \phi_i(L'(i), L(-i))$. This completes the proof of truthfulness.

Now we proceed to prove the Pareto optimality of ϕ . Suppose there exists another allocation ϕ' such that ϕ' is a Pareto improvement of ϕ , e.g. $\phi'_i(L(i), L(-i)) \succeq_i \phi_i(L(i), L(-i))$, $\forall i \in [n]$ and there exists $k \in [n]$ such that $\phi'_k(L(k), L(-k)) \succ_k \phi_k(L(k), L(-k))$. Let $u = \min\{k : \phi'_k(L(k), L(-k)) \succ_k \phi_k(L(k), L(-k))\}$. We know $\phi'_i(L(i), L(-i)) \simeq_i \phi_i(L(i), L(-i))$, $\forall i < u$. Suppose $a = \phi'_u(L(u), L(-u))$ and $b = \phi_u(L(u), L(-u))$ and $a \succ_u b$, we have the following claim:

Claim 4.5. *There is a path from a to a labeled agent in $TG(u, a, [u-1])$ or $a \in A_u$ at the beginning of u th loop of for loop.*

Proof. Suppose $a \notin A_u$, which means that a has been allocated to some agent $i \leq u$ in SDMT-2. Consider the graph $TG(u, a, [u-1])$, Suppose, a is allocated to $u_1 < u$, then $a \simeq \phi_{u_1}(L(u_1), L(-u_1)) \simeq \phi'_{u_1}(L(u_1), L(-u_1))$, if u_1 owns more than one object, e.g. owning a and $\phi'_{u_1}(L(u_1), L(-u_1))$, then we are done. Otherwise, by Observation 4.2, $\phi'_{u_1}(L(u_1), L(-u_1))$ must be allocated to some agent $u_2 < u$ in $TG(u, a, [u-1])$, different from u_1 (if not different, we are done), by the same argument that check whether u_2 simultaneously owns $\phi'_{u_1}(L(u_1), L(-u_1))$ and $\phi'_{u_2}(L(u_2), L(-u_2))$, if not then $\phi'_{u_2}(L(u_2), L(-u_2))$ must be allocated to some agent $u_3 < u$, different from u_1, u_2 , (otherwise we are done). Continue this procedure, since there are only finite agents ($u-1$ agents), we must finally arrive at an agent who owns more than one object and get a path from a to this agent. \square

Clearly, by Claim 4.5, we know agent u will be allocated to an object no inferior to a , e.g. $b \succeq_u a$, contradiction.

We finally establish the running time of SDMT-2. By previous analysis, in each loop i , the running time is $O(|L(i)| + (i-1)\gamma)$, sum i over $[n_1]$, we obtain that the running time of SDMT-2 is $O(mn + n^2\gamma)$. However, we can reduce the running time of *Trading*($i, \max TG(i, L(i), [i-1]), [i-1]$) to $O(i + \gamma + (i-1)\gamma)$ by analysing process carefully. When constructing $TG(i, L(i), [i-1])$, we need only to consider all the objects in $\cup_{k=1}^j C_k^i$ such that $|\cup_{k=1}^j C_k^i| \geq i$, that is only constructing $TG(i, \cup_{k=1}^j C_k^i, [i-1])$. This also reduces the complexity of finding j_2 to $O(i + \gamma + (i-1)\gamma)$. Similarly, when defining j_1 , we need only to consider whether $A_i \cap (\cup_{k=1}^j C_k^i)$ is empty set or not, where $|\cup_{k=1}^j C_k^i| \geq i$. This is modified SDMT-2. The reason is as follows: If i is allocated some objects in SDMT-2, then all these objects must be in $\cup_{k=1}^j C_k^i$. This is because we will have either $A_i \cap (\cup_{k=1}^j C_k^i) \neq \emptyset$ or there is a path from an object in $\cup_{k=1}^j C_k^i$ to a labeled agent, where $|\cup_{k=1}^j C_k^i| \geq i$ (if $A_i \cap (\cup_{k=1}^j C_k^i) = \emptyset$, then all elements in $\cup_{k=1}^j C_k^i$ are allocated to $[i-1]$, since $|\cup_{k=1}^j C_k^i| \geq i$, then there exists an agent owning more than one object, hence, there

Algorithm 3: Random SDMT-2 for Weighted Agents with Ties

Input: Agents N ; Objects A ; Preference list profile L ; Weights W

Output: Matching

for each agent $i \in N$ **do**

 | Pick $Y_i \in [0, 1]$ uniformly at random;

end

Sort agents in decreasing order of $w_i(1 - g(Y_i))$ (break ties in favor of smaller index);

Run SDMT-2 according to above order;

Return the matching;

is a path (edge) from an object in $\cup_{k=1}^j C_k^i$ to this labeled agent). If i is allocated nothing in SDMT-2, then we know the modified SDMT-2 will allocate i nothing. Hence, the total running time of modified SDMT-2 is $\sum_i O(i + \gamma + (i - 1)\gamma) = O(n^2\gamma)$. \square

4.3 Randomized mechanism

We now present a universally truthful and Pareto optimal mechanism with approximation ratio of $\frac{e}{e-1}$, where agents may have weights and their preferences may involve ties (see Algorithm 3, where $g(y) = e^{y-1}$). When preferences are strict, Algorithm 3 reduces to a variant of RSDM that has been used in weighted online bipartite matching with approximation ratio $\frac{e}{e-1}$ (see [3] and [15]). Our analysis of Algorithm 3 is a non-trivial extension of the primal-dual analysis from [15] to the case where agents' preferences may involve ties.

To gain some high-level intuition behind this extension, we highlight here the similarities and differences between our problem and that of online bipartite matching. Our problem with strict preferences and without weights is closely related to online bipartite matching.⁶ If each agent in our problem ranks his desired objects in the order that precisely follows the arrival order of objects in the online bipartite matching, the two problems are equivalent. Therefore, we extend the analysis of this particular setting, where each agent's preference list is a sublist of a global preference list, to the general case where agents preferences are not constrained and may involve ties, and furthermore agents may have weights.

If the weights are public, Algorithm 3 is universally truthful and Pareto optimal. This is because it chooses a random order of the agents, given the weights, and then runs SDMT-2 according to this order. It follows from SDMT-2 that, if the order of the other agents is given, an agent can get a better object if he appears earlier in this order. Then it is not difficult to see that if the weights are private, and under the assumption that no agent is allowed to bid over his private weight, Algorithm 3 is still universally truthful in the sense that no agent will lie about his preferences or weight.

Theorem 4.6. *Algorithm 3 is universally truthful, even if the weights and preference lists of the agents are their private knowledge, assuming that no agent can over-bid his weight.*

⁶In the online bipartite matching problem [8], vertices of one partition (think of them as agents) are given and fixed, while vertices of the other partition (think of them as objects) arrive in an adversarial order. When an item arrives, we get to see the incident edges on agents. These edges indicate the set of agents that desire this object. The algorithm must immediately match this object to one of the unmatched agents desiring it (or choose to throw it away). In the end, the size of the obtained matching is compared with the optimum matching in the realized graph.

$$\begin{aligned}
& \max \sum_{(i,a) \in E} w_i x_{ia} \quad s.t. \quad \forall i \in N : \sum_{a:(i,a) \in E} x_{ia} \leq 1; \quad \forall a \in A : \sum_{i:(i,a) \in E} x_{ia} \leq 1; \\
& \quad \quad \quad \forall (i,a) \in E : x_{ia} \geq 0. \\
& \min \sum_{i \in N} \alpha_i + \sum_{a \in A} \beta_a \quad s.t. \quad \forall (i,a) \in E : \alpha_i + \beta_a \geq w_i; \quad \forall i \in N, a \in A : \alpha_i, \beta_a \geq 0.
\end{aligned}$$

Figure 2: The primal and dual LP relaxation. $G = (V, E)$, where $V = N \cup A$ and $E = \{(i, a), i \in N, a \in A\}$

Proof. Algorithm 3 is a distribution of deterministic mechanisms due to the selection of random variables Y_i . For each deterministic mechanism (i.e. SDMT-2 when $Y_i, i \in N$ is fixed), we prove it is truthful w.r.t. weights and preference lists. Denote by ϕ the mechanism of SDMT-2 when $Y_i, i \in N$ is fixed. If is not difficult to see that for any (W, L) , $w'_i \leq w_i$ and $L'(-i), i \in N$, we have $\phi_i(W, L) \succeq_i \phi_i((w'_i, w_{-i}), L) \succeq_i \phi_i((w'_i, w_{-i}), (L'(i), L(-i)))$. The first preferred order in this chain follows from the fact that the order of i when i bids w_i is better than or equal to his order when he bids w'_i . The second preferred order in this chain follows by the truthfulness of SDMT-2 when weights are public. \square

5 Analysis of the approximation ratio

To analyze the approximation ratio of Algorithm 3, we first write the LP formulation of the (relaxed) problem and its dual LP formulation. Given random variables Y_i , we will define a primal solution and a dual solution obtained by Algorithm 3, which are both random variables, such that the objective value of the primal solution is always at least a fraction F of the objective value of the dual solution, and that the expectation of duals is feasible. Hence, the expectation of the primal solution is at least F times the expectation of duals, which by weak LP duality, is at least optimal value of the primal LP. The standard LP and its dual of our problem are given in Figure 2.

By Lemma 5.1, proved in [15], the inverse of approximation ratio is $F \in [0, 1]$.

Lemma 5.1 ([15]). *Suppose that a randomized primal-dual algorithm has a primal feasible solution with value P (which is a random variable) and a dual solution which is not necessarily feasible, with value D (which is also a random variable) such that*

1. *for some universal constant F , $P \geq F \cdot D$, always, and*
2. *the expectation of the randomized dual variables form a feasible dual solution.*

The expectation of P is then at least $F \cdot \text{OPT}$ where OPT is the value of the optimum solution.

Proof. Since $P \geq F \cdot D$, taking expectations, $\mathbb{E}(P) \geq F \cdot \mathbb{E}(D)$. The cost of the dual solution obtained by taking the expectations of the randomized dual variables is $\mathbb{E}(D)$ and they form a feasible dual solution, therefore $\mathbb{E}(D) \geq \text{OPT}$. Hence, $\mathbb{E}(P) \geq F \cdot \text{OPT}$. \square

Note that in Lemma 5.1, OPT is the value of maximum weighted matching, which is no less than the value of maximum weighted Pareto optimal matching. Hence, if the condition of Lemma 5.1 holds, the approximation ratio of the mechanism is $\frac{1}{F}$. The construction of the duals depends on function g . Let $F = (1 - \frac{1}{e})$. It is not difficult to see that

$$\forall \theta \in [0, 1] \int_0^\theta g(y) dy + 1 - g(\theta) \geq F \quad (1)$$

For any random selection of $Y_i, i \in N$, let $\vec{Y} = (Y_1, Y_2, \dots, Y_{n_1}) = (Y_i, Y_{-i})$. Following the procedure of Algorithm 3, whenever agent i is matched to object a , let

$$x_{ia}(\vec{Y}) = 1, \quad \alpha_i(\vec{Y}) = w_i g(Y_i)/F, \quad \beta_a(\vec{Y}) = w_i(1 - g(Y_i))/F.$$

For all unmatched i and a , set $x_{ia}(\vec{Y}) = \alpha_i(\vec{Y}) = \beta_a(\vec{Y}) = 0$. By this definition, it follows that for any $Y_i, i \in N$, the random value P of primal solution $\{x_{ia}(\vec{Y}), i \in N, a \in A\}$ is always at least $F \cdot D$, where D is random value of the dual solution $\{\alpha_i(\vec{Y}), i \in N, \beta_a(\vec{Y}), a \in A\}$.

Hence, to satisfy the conditions of Lemma 5.1, we need to show that the expectation of dual solution $\{\alpha_i(\vec{Y}), i \in N, \beta_a(\vec{Y}), a \in A\}$ is feasible for the dual LP, implying that the approximation ratio of Algorithm 3 is at most $\frac{1}{F} = \frac{e}{e-1}$. The main technical difficulty lies in proving the dominance lemma and the monotonicity lemma (see Lemma 5.3 and 5.4). To prove these two lemmas, we define a threshold which specifies whether an agent still gets matched if he is put back into the order after he is first removed from it. For an agent with strict preferences, such a threshold is the same as that of defined in the online bipartite matching problem. However, in the presence of ties, the same defined threshold does not work. We show how to define such a threshold for our algorithm.

Fix an agent $i \in [n_1]$ and object $a \in A$, such that $(i, a) \in E$. Fix Y_{-i} , the random variables $Y_{i'}$ for all other agents $i' \neq i$. We use σ to denote the order of agents under Y_{-i} , e.g., $\sigma(1)$ is the first agent, etc, and $\sigma([i]) = \{\sigma(1), \sigma(2), \dots, \sigma(i)\}$. Consider Algorithm 3 running on the instance without agent i and let us denote this procedure by ALG_{-i} , where σ is the order of agents under ALG_{-i} . The threshold y^c is then defined as follows:

1. If a is unmatched in ALG_{-i} , let $y^c = 1$.
2. Otherwise, suppose that a is matched in ALG_{-i} to some agent i' . Then consider the allocations just after the for loop in ALG_{-i} terminated.
If i' is labeled, set $y^c = 1$.
3. Otherwise, suppose $a \in C_j^{i'}$ and construct the trading graph $TG(i', C_j^{i'} \setminus \{a\}, [n_1] \setminus \{i\})$ from all the objects in $C_j^{i'}$ other than a (note that $\sigma([n_1 - 1]) = [n_1] \setminus \{i\}$).
If there is a path in $TG(i', C_j^{i'} \setminus \{a\}, [n_1] \setminus \{i\})$ from i' to a labeled agent, set $y^c = 1$.
4. Otherwise, define
 $i'' = \arg \min \{w_l(1 - g(Y_l)) \mid \text{there is a path from } i' \text{ to } l \text{ in } TG(i', C_j^{i'} \setminus \{a\}, [n_1] \setminus \{i\})\}$.
If $w_i(1 - g(y)) = w_{i''}(1 - g(Y_{i''}))$ has a solution $y \in [0, 1]$ define y^c to be this solution.
($g(y)$ is strictly increasing so if there is a solution, it is unique)
5. Otherwise define y^c to be 0.

Now consider Algorithm 3 running on the original instance (denote such procedure as ALG), with (Y_i, Y_{-i}) fixed. Suppose that τ is the order of agents under this execution of ALG . The intuition behind the definition of y^c is the following. Having Y_{-i} fixed, we want to define y^c such that if we run ALG with (Y_i, Y_{-i}) where $Y_i < y^c$, then agent i gets matched (to any of his objects). If (1) holds, then $Y_i < 1$ and i will be matched because at least object a is his available candidate. If (2) happens, then $Y_i < 1$ and i will also be matched because object a can be re-allocated from the labeled agent i' to i . Case (3) is analogous to (2) with the only difference that we now have a trading path from i' to a labeled agent. Case (4) will be discussed just after Observation 5.2.

Observation 5.2. Assume that agent i is unmatched in his turn in the “for loop” of ALG . Suppose $\tau(u) = i$, which means i selects his object in u th iteration of the “for loop”. Then at end of k th iteration of the “for loop”, $k \geq u$, there is no path from i to a labeled agent in $TG(i, L(i), \tau([k]))$ (meaning this graph is frozen).

Proof. By SDMT-2, we know $A_u \cap L(\tau(u)) = A_u \cap L(i) = \emptyset$, which means all the objects in $L(i)$ have been allocated to agents $\tau([u-1])$. Since $\tau(u)$ is unmatched, there is no path from $\tau(u)$ to any labeled agent in $TG(\tau(u), L(\tau(u)), \tau([u-1]))$. Let $S \subseteq \tau([u-1])$ be the set reachable from $\tau(u)$ in $TG(\tau(u), L(\tau(u)), \tau([u-1]))$, clearly, each agent in S is unlabeled. Actually any agent in S is frozen. Therefore, any path through i after u th iteration will reach an unlabeled agent. \square

The following two properties (dominance and monotonicity) are well known for agents with strict preference orderings. We generalize them to agents with indifference. The difficulty of proving both dominance and monotonicity lemmas (Lemma 5.3 and 5.4) lies in case (4). This is our main technical contribution as compared to the analysis in [15].

Recall that τ (σ , resp.) is the order of agents under the execution of ALG (ALG_{-i} , resp.). We first deal with the case (4) of the dominance lemma. Note that in this case there is a path from i' to i'' in $TG(i', C_j^{i'} \setminus \{a\}, [n_1] \setminus \{i\})$ and agent i'' is unlabeled. We will prove dominance by contradiction, using the following two main steps. Suppose $\sigma(u) = i''$ in ALG_{-i} , then $\tau(u+1) = i''$ in ALG under case (4). Suppose that $Y_i < y^c$ and i is not matched in ALG , then the outcome of ALG is the same as that of ALG_{-i} for all the other agents except i . Based on this fact, first, we prove that either i' is labeled or there is a path, call it P_1 , from i' to a labeled agent in $TG(i', C_j^{i'}, \tau([u]))$ at the end of the u th iteration of the for loop in ALG . Second, due to the above property, we argue that there is a path, say P_2 , from i to a labeled agent in $TG(i, a, \tau([u]))$ at the end of the u th iteration of the for loop in ALG , contradicting Observation 5.2; thus i will be matched. Path P_2 is constructed by the concatenation of arc (i, i') and path P_1 , or the concatenation of arc (i, i''') , for some i''' on path P_1 , and the rest of path P_1 . The existence of P_1 is proved by a careful analysis of the structure of frozen subgraphs of the trading graph as the algorithm proceeds; the details can be found in the proof of Lemma 5.3.

Lemma 5.3 (DOMINANCE LEMMA). *Given Y_{-i} , i gets matched if $Y_i < y^c$.*

Proof. Case 1: If a is unmatched in ALG_{-i} , then $y^c = 1$. Suppose agent i is unmatched in ALG , and procedure ALG is the same as ALG_{-i} for all the other agents except i . Then a is always available to agent i , meaning a will be matched to agent i by process of SDMT-2, contradiction.

Case 2: If a is matched to i' in ALG_{-i} :

(i): If i' is labeled or if there is a path from i' to a labeled agent in $TG(i', C_j^{i'} \setminus \{a\}, [n_1] \setminus \{i\})$, if i is unmatched, then there is a path from i to a labeled agent in $TG(i, a, [n_1])$, then by $Trading(i, a, [n_1])$, we get a Pareto improvement, contradicting that SDMT-2 is Pareto optimal.

(ii): The case $y^c = 0$ is trivial, so we consider that $w_i(1 - g(y)) = w_{i''}(1 - g(Y_{i''}))$ has a solution. Suppose that $\sigma(u) = i''$ in ALG_{-i} , then if $Y_i < y^c$, we know $w_i(1 - g(Y_i)) > w_{i''}(1 - g(Y_{i''}))$ meaning the agent i is prior to agent i'' in ALG . Then $\tau(u+1) = i''$ in ALG . If i is unmatched in ALG , then procedure ALG is the same as ALG_{-i} for all the other agents except i .

Suppose i'' is allocated an object b in ALG . If $i'' = i'$, then $b = a$, and if in addition $a \in A_{u+1}$, this means a is always available to all the agents prior to $\tau(u+1) = i'' = i'$. Therefore, a will be available to i when i initially selects objects, implying that i must be allocated to some objects in his turn, contradiction.

The case $a \notin A_{u+1}$ is same as the case $i'' \neq i'$, so we consider that $i'' \neq i'$. Since there is a path from i' to i'' after the for loop in ALG terminates, due to our assumption i is still unmatched. Suppose in this path $\tau(k)$ points to $\tau(u+1) = i''$, for some $k \leq u$, then b is available to $\tau(k)$ or b has been allocated before k th iteration of the for loop in ALG . Since finally $\tau(k)$ gets an object in the same indifference class as b by Observation 4.1,

before $(u + 1)$ st iteration of the for loop in ALG , b has been allocated by Observation 4.2. Hence, in the $(u + 1)$ st iteration of for loop in ALG , $\tau(u + 1)$ gets object b through the trading graph. Observe that the trading graph $TG(i', C_j^{i'} \setminus \{a\}, [n_1])$ after the for loop in ALG terminates is exactly the same as $TG(i', C_j^{i'} \setminus \{a\}, \tau([u + 1]))$ at the end of $(u + 1)$ st iteration (otherwise, some agent $\tau(\ell)$ may be reachable from i' , where $\ell > u + 1$, by process of SDMT-2, contradicting to the definition of i'').

Therefore, at the end of u th iteration of the for loop of ALG , suppose that B is the set of objects allocated to i' . Then we have the following three cases (noting ALG is the same as ALG_{-i}) for all the other agents except i :

Case (ii-1): i' is labeled, then $a \in B$. Otherwise, if $a \notin B$, then in $(u + 1)$ st for loop iteration of ALG , a will not be allocated to i' at the end this $(u + 1)$ st iteration by the process of SDMT-2. Thus $a \in B$, and since the trading graph $TG(i', C_j^{i'} \setminus \{a\}, [n_1])$ after the for loop in ALG terminates is exactly the same as $TG(i', C_j^{i'} \setminus \{a\}, \tau([u + 1]))$ at the end of $(u + 1)$ st iteration, then i' will not be matched to a at the end of for loop of ALG , contradiction.

Case (ii-2): i' is unlabeled and $B = \{a\}$, then there is a path from i' to a labeled agent in $TG(i', C_j^{i'} \setminus \{a\}, \tau([u]))$. Otherwise, all the agents reachable from i' are frozen after u th iteration of the for loop, meaning their allocations are fixed, since all the objects in their indifference class where they are allocated to some objects have been allocated by Observation 4.2. Thus, $TG(i', C_j^{i'} \setminus \{a\}, \tau([u]))$ should be the same as $TG(i', C_j^{i'} \setminus \{a\}, \tau([u + 1]))$, however, since $\tau(u + 1)$ is reachable from i' in $TG(i', C_j^{i'} \setminus \{a\}, \tau([u + 1]))$ while $\tau(u + 1)$ does not appear in $TG(i', C_j^{i'} \setminus \{a\}, \tau([u]))$, contradiction.

Case (ii-3): i' is unlabeled and $B = \{c\}$, where $c \neq a$. Then there is a path from i' to a labeled agent in $TG(i', \{a\}, \tau([u]))$. Otherwise a and the agent matched to a is frozen at the end of u th for loop iteration in ALG , which means a will not matched to i' at the end of for loop of ALG , contradiction.

As a result, in either of the above three cases, there is a path from i to a labeled agent in $TG(i, a, \tau([u]))$ at the end of u th for loop iteration in ALG (for case (ii-1), i points to i' , which is labeled in $TG(i, a, \tau([u]))$; for case (ii-2), i points to i' in $TG(i, a, \tau([u]))$ and there is a path from i' to a labeled agent in $TG(i', C_j^{i'} \setminus \{a\}, \tau([u])) \subseteq TG(i, a, \tau([u]))$; finally, for case (ii-3), suppose a is assigned to i''' at the end of u th for loop iteration, then there is a path from i''' to a labeled agent in $TG(i, a, \tau([u]))$ and i points to i''' in $TG(i, a, \tau([u]))$). This contradicts Observation 5.2. Hence, i must be matched. \square

Let $\beta_a(\vec{Y})$ be β_a^s when ALG is the procedure of Algorithm 3 running on the original instance when Y_{-i} is fixed and $Y_i = s$, i.e., $\beta_a^s = \beta_a((s, Y_{-i}))$. Let $\beta_a^c = \beta_a^{y^c} = w_i(1 - g(y^c))/F$.

The difficulty of the proof of the monotonicity lemma (Lemma 5.4) still lies in case (4). We prove it in three steps. Recall that τ (σ , respectively) is the order of agents under the execution of ALG (ALG_{-i} , respectively). Let $\sigma(u) = i''$ in ALG_{-i} . Observe that the monotonicity lemma means that a is allocated to an agent prior to i'' or i''' . The proof of this is easy in the case where $Y_i > y^c$. To see this, note that ALG and ALG_{-i} result in the same tentative allocation at the end of their u th loop, since i is inserted back after i'' . Hence, we only need to consider the case where $Y_i < y^c$, which implies that i is inserted back prior to i'' . First, we prove that no agent, except i , is allocated a better object in ALG compared to ALG_{-i} . The argument is by contradiction: suppose there exists an agent i''' who receives a better object in ALG than in ALG_{-i} . Then i must be inserted

before i''' . Consequently, there exists an agent s prior to i''' who will get a worse object in ALG than in ALG_{-i} . Based on this fact, and using an alternating path argument, it is proved that there exists a path from s to i''' in s 's trading graph constructed from a higher indifference class of s (than s 's allocated indifference class in ALG) after i''' is allocated in ALG . This contradicts the fact that this path should not exist since the graph from that higher indifference class is frozen. Secondly, if i' gets a worse object in ALG compared to ALG_{-i} , we prove that a must be allocated to an agent prior to i' , which is in turn prior to i'' . The reason is as follows: by Observation 4.2, a must be allocated and frozen before i' is considered in ALG . Thirdly, if i' gets an object in ALG in the same indifference class as a , then we prove that there exists an agent s^* prior to i'' , and suppose $\tau(u^*) = s^*$, such that there is a path from s^* to i' in $TG(s^*, C_{j^*}^{s^*}, \tau([u^*]))$ at the end of the u^* th loop of ALG , where $C_{j^*}^{s^*}$ is the indifference class in which s^* is allocated an object in ALG_{-i} . As a consequence, by Observation 4.2, a is allocated to an agent prior to s^* and all the agents reachable from s^* in $TG(s^*, C_{j^*}^{s^*}, \tau([u^*]))$ are frozen, then a will be finally allocated to an agent prior to s^* in ALG . This means that a is allocated to an agent prior to i'' . This gives the monotonicity lemma, which together with dominance is used to prove Lemma 5.7.

Lemma 5.4 (MONOTONICITY LEMMA). *Given Y_{-i} , for all choices of Y_i , $\beta_a^{Y_i} \geq \beta_a^c$.*

Proof. **Case 1:** If a is unmatched in ALG_{-i} , or if a is matched to i' in ALG_{-i} and i' is labeled, or a is matched to i' in ALG_{-i} and there is a path from i' to a labeled agent in $TG(i', C_j^{i'} \setminus \{a\}, [n_1])$, then $y^c = 1$ and $\beta_a^c = w_i(1 - g(y^c))/F = 0$, so $\beta_a^{Y_i} \geq \beta_a^c = 0$.

Case 2: If a is matched to i' in ALG_{-i} , there is no path from i' to a labeled agent in $TG(i', C_j^{i'} \setminus \{a\}, [n_1] \setminus \{i\})$. Suppose $\sigma(u) = i''$ in ALG_{-i} . By previous analysis, the trading graph $TG(i', C_j^{i'} \setminus \{a\}, \sigma([u]))$ at the end of u th loop is the same as $TG(i', C_j^{i'} \setminus \{a\}, [n_1] \setminus \{i\})$ at the end of for loop in ALG_{-i} (otherwise, the $TG(i', C_j^{i'} \setminus \{a\}, \sigma([u]))$ is not frozen after u th loop of ALG_{-i} , meaning there is a path from i' to a labeled agent in $TG(i', C_j^{i'} \setminus \{a\}, \sigma([u]))$. Therefore, either i' will reach an agent inferior to i'' or a labeled agent in $TG(i', C_j^{i'} \setminus \{a\}, [n_1] \setminus \{i\})$ by SDMT-2. This contradicts the definition of i'').

Suppose that equation $w_i(1 - g(y)) = w_{i''}(1 - g(Y_{i''}))$ does not have a solution, which means that $y^c = 0$ and $w_i(1 - g(Y_i))/F < w_{i''}(1 - g(Y_{i''}))/F$, for any $Y_i \in [0, 1]$. This shows that the process is the same for agents prior to agent i'' until the end of u th loop in ALG and ALG_{-i} . Since there is no path from i' to a labeled agent in $TG(i', C_j^{i'} \setminus \{a\}, \sigma([u]))$, the agents reachable from i' are frozen. Hence, a will be finally still allocated to i' in ALG , implying $\beta_a^{Y_i} = w_{i'}(1 - g(Y_{i'}))/F \geq w_{i''}(1 - g(Y_{i''}))/F > \beta_a^c = w_i(1 - g(0))/F$.

Now consider the last case that equation $w_i(1 - g(y)) = w_{i''}(1 - g(Y_{i''}))$ has a solution, then $\beta_a^c = w_i(1 - g(y^c))/F = w_{i''}(1 - g(Y_{i''}))/F$. Consider the following three cases:

Case (2-i): If $Y_i > y^c$, this means $w_i(1 - g(Y_i))/F < w_{i''}(1 - g(Y_{i''}))/F$, the analysis of this case is the same as above (the case that equation $w_i(1 - g(y)) = w_{i''}(1 - g(Y_{i''}))$ does not have a solution) since i will select objects after i'' . Thus, we have $\beta_a^{Y_i} = w_{i'}(1 - g(Y_{i'}))/F \geq \beta_a^{y^c} = \beta_a^c$.

Case (2-ii): If $Y_i < y^c$, then $w_i(1 - g(Y_i))/F > w_{i''}(1 - g(Y_{i''}))/F$, which means that i is prior to i'' in ALG . We have the following claim:

Claim 5.5. *No agent can get a better object in ALG than in ALG_{-i} after inserting i into some position from 1 to u .*

Proof. Suppose otherwise suppose there exists an agent getting a better object, let k be the smallest position where such agents are placed in ALG . Then i must be inserted in a position before k (otherwise, the process is the same for the first k agents in ALG_{-i} and ALG , so agent $\tau(k)$ can not get a better object). Let $i''' = \tau(k)$. Suppose i''' gets object b

in ALG and c in ALG_{-i} , where $b \succ_{i'''} c$. Observe that $\sigma(k-1) = i'''$ in ALG_{-i} . Suppose $b \in C_j^{i'''}$. Consider the trading graph $TG(i''', C_j^{i'''}, \sigma([k-1]))$ at the end of $(k-1)$ st loop of ALG_{-i} .

Let S be the agents reachable from i''' in $TG(i''', C_j^{i'''}, \sigma([k-1]))$ at the end of $(k-1)$ st loop of ALG_{-i} . Note that any agent in S is prior to i''' . Any agent in S is allocated only one object and frozen in ALG_{-i} . Since in ALG , b is allocated to i''' , then in k th loop of ALG , i''' will be allocated some objects in $C_j^{i'''}$, which means some agent in S will get worse compared to the allocation in ALG_{-i} .

The reason is as follows: No one can get better by the definition of k . If all the agents in S can remain the same in ALG compared with ALG_{-i} (i.e., get the objects in the same indifference class in ALG and in ALG_{-i}). The only possible allocation of S in ALG is reallocating all the objects matched to S in ALG_{-i} to S again such that each agent gets exactly one object. If there is some extra object e in ALG allocated to agent S in ALG , then e must be allocated to some agent j in ALG_{-i} , since e in ALG allocated to some agent in S , thus, j can be reached by some agent in S in ALG_{-i} . Thus, $j \in S$, a contradiction. All the objects in $C_j^{i'''}$ has been allocated to some agents in S . In ALG , we will need to allocate $|S|$ objects to $S \cup \{i'''\}$ agents because some objects owned by S in ALG_{-i} will be allocated to agent i''' . This is not possible, a contradiction.

Let s be an agent in S who gets worse and there is a path from s to i''' in trading graph $TG(s, d, \tau[k])$ at the end of k th loop in ALG , where d is the allocated object of s in ALG_{-i} . (Such an agent must exist: it can be found by the following procedure. Suppose $d_1 \simeq_{s_1} b$ owned by s_1 in ALG_{-i} is allocated to i''' in ALG at the end of k th loop in ALG . If s_1 gets worse in ALG compared to ALG_{-i} , then s_1 is an agent we are looking for. Otherwise, s_1 will be allocated d_2 owned by $s_2 \in S$ in ALG_{-i} at the end of k th loop of ALG . If s_2 gets a worse object, then s_2 is an agent we are looking for. Otherwise, continue this procedure. By finiteness of S and the fact that agents in S own $|S|$ objects in ALG_{-i} , these objects will be allocated to agents in $S \cup \{i'''\}$ in ALG , and one of these objects will be allocated to i''' . Thus, we can find such an agent. The path from s to i''' in trading graph $TG(s, d, \tau[k])$ at the end of k th loop in ALG is just the reverse path by the above procedure). Suppose $\tau(\ell) = s$ in ALG and $d \in C_h^s$, consider ℓ th loop in ALG , all the agents reachable from s in $TG(s, C_h^s, \sigma([\ell]))$ are frozen and prior to agent s since s does not obtain any object in the indifference class C_h^s .

This contradicts the fact that there is a path from s to i''' (which is inferior to s) in trading graph $TG(s, d, \tau[k])$ at the end of k th loop in ALG . \square

Claim 5.6. a must be allocated to an agent prior to i'' or to i'' .

Proof. Suppose $\sigma(u_1) = i'$ and $\sigma(u) = i''$ in ALG_{-i} . The following two cases are considered:

Case (1). If i' gets worse, meaning he gets worse object in ALG than a in ALG_{-i} , then $\tau(u_1 + 1) = i'$ in ALG (i is inserted back prior to i'). Thus, all the agents reachable from i' in $TG(i', a, \tau([u_1 + 1]))$ are frozen and the agent who owns a will finally get a , this agent is prior to i' , giving that $\beta_a^{Y_i} \geq w_{i'}(1 - g(Y_{i'})) \geq w_{i''}(1 - g(Y_{i''})) = \beta_a^c$.

Case (2). If i' gets a in ALG , then we are done. Otherwise, suppose i' gets an object $a' \simeq_{i'} a$, $a', a \in C_j^{i'}$ in ALG . Denote by S^* the set of agents reachable from i' in $TG(i', C_j^{i'} \setminus \{a\}, \sigma([n_1 - 1]))$ at the end of for loop of ALG_{-i} (note that $\sigma([n_1 - 1]) = [n_1] \setminus \{i\}$). If no one in S^* gets worse in ALG than in ALG_{-i} , then a must be allocated to some agent in S^* . The reason is similar to the above argument. All agents in S^* get exactly one object, if a is not allocated in S^* , no one gets worse in S^* , there must be an extra object b allocate to some agent j in S^* . No matter whom b is allocated to in ALG_{-i} , there is a path from j to this agent, hence this agent belongs to S^* , a contradiction.

Note that, by the definition of i'' , for any $s \in S^*$, $\sigma^{-1}(s) > \sigma^{-1}(i'')$ ($\sigma^{-1}(s)$ denotes the order of s in σ or in ALG_{-i}) implies that $w_s(1 - g(Y_s)) \geq w_{i''}(1 - g(Y_{i''}))$. Therefore $\beta_a^{Y_i} \geq w_{i''}(1 - g(Y_{i''})) = \beta_a^c$.

Otherwise, by the previous argument, there exists $s^* \in S^*$ who gets worse in ALG compared to ALG_{-i} , and there is a path from $s^* \in S^*$ to i' in $TG(s^*, d^*, [n_1])$ at the end of ALG , where d^* is the allocation of s^* in ALG_{-i} . If s^* is prior to i' , then by similar argument as above, there should be no path from s^* to an agent inferior to s^* (constructed from the objects in $L(s^*)$ no worse than d^*) at the end of ALG , a contradiction. Hence, s^* can only be inferior to i' . Suppose $\tau(u^*) = s^*$ and $d^* \in C_{j^*}^{s^*}$, then we know there is a path from s^* to i' in $TG(s^*, C_{j^*}^{s^*}, [n_1])$ at the end of for loop of ALG . Next we will prove the following:

There is a path from s^* to i' in $TG(s^*, C_{j^*}^{s^*}, \tau([u^*]))$ at the end of u^* th loop of ALG . (*)

If (*) is true, then all the agents reachable from s^* in $TG(s^*, C_{j^*}^{s^*}, \tau([u^*]))$ are frozen, and a is allocated to an agent prior to s^* due to Observation 4.2. Then, we have $\beta_a^{Y_i} \geq w_{s^*}(1 - g(Y_{s^*})) \geq w_{i''}(1 - g(Y_{i''})) = \beta_a^c$ since $s^* \in S^*$. Suppose (*) is not true, then all the agents U^* reachable from s^* in $TG(s^*, d^*, \tau([u^*]))$ have been frozen. U^* will remain the same until the end of ALG and $i' \notin U^*$. However, by the definition of S^* and by $s^* \in S^*$, there is a path from s^* to i' in $TG(s^*, d^*, [n_1])$ at the end of ALG , meaning $i' \in U^*$, a contradiction. \square

By Claim 5.6, we know that if $Y_i < y^c$, $\beta_a^{Y_i} \geq w_{i''}(1 - g(Y_{i''}))/F = \beta_a^c$.

Case (2-iii): If $Y_i = y^c$, this means $w_i(1 - g(Y_i))/F = w_{i''}(1 - g(Y_{i''}))/F$. If $i'' < i$, the case is same as $Y_i > y^c$. Otherwise, it falls into the case $Y_i < y^c$.

In all, for all choices of Y_i , $\beta_a^{Y_i} \geq \beta_a^c$. \square

Lemma 5.7 ([15]). $\forall (i, a) \in E$, $\mathbb{E}_{\vec{Y}}(\alpha_i(\vec{Y}) + \beta_a(\vec{Y})) \geq w_i$.

Proof. Fixed choices of Y_{-i} , by Dominance Lemma (Lemma 5.3), i is matched whenever $Y_i < y^c$. Hence,

$$\mathbb{E}_{Y_i}(\alpha_i(\vec{Y})) \geq w_i \int_0^{y^c} g(y) dy / F.$$

By the Monotonicity Lemma (Lemma 5.4), $\beta_a(\vec{Y}) = \beta_a^{Y_i} \geq \beta_a^c = w_i(1 - g(y^c))/F$, for any $Y_i \in [0, 1]$, then

$$\mathbb{E}_{Y_i}(\beta_a(\vec{Y})) \geq w_i(1 - g(y^c))/F.$$

Therefore, note that by formula (1), we have

$$\mathbb{E}_{Y_i}(\alpha_i(\vec{Y}) + \beta_a(\vec{Y})) \geq w_i \int_0^{y^c} g(y) dy / F + w_i(1 - g(y^c))/F \geq w_i.$$

As a result, $\mathbb{E}_{\vec{Y}}(\alpha_i(\vec{Y}) + \beta_a(\vec{Y})) \geq w_i$. \square

From Lemma 5.1 and Lemma 5.7, we have the following theorem.

Theorem 5.8. *Algorithm 3 achieves approximation ratio of $\frac{e}{e-1}$ for weighted agents with indifference.*

6 Lower bounds

Preliminaries. We will use Yao’s minmax principle [33] to obtain a non-trivial lower bound for universally truthful and Pareto optimal mechanisms and another lower bound for an “online” version of our problem. We first need some preliminaries.

Let us fix the number of agents n_1 and the number of objects n_2 . The number of distinct instances and the number of deterministic truthful and Pareto optimal mechanisms are finite. Denote by \mathcal{T} the set of deterministic truthful and Pareto optimal mechanisms with input size n_1 and n_2 , and \mathcal{I} the set of instances with input size n_1 and n_2 . Let \mathbf{P} and \mathbf{Q} denote the set of probability distributions on \mathcal{T} and \mathcal{I} , respectively. Denote $\mathbb{E}_{p,q}(r(T_p, I_q))$ as the inverse of approximation ratio when the input distribution is $q \in \mathbf{Q}$ on \mathcal{I} and a universally truthful mechanism and Pareto optimal mechanism T_p taken from \mathcal{T} with probability $p \in \mathbf{P}$. Then the minmax theorem [33] states the following:

$$\min_{q \in \mathbf{Q}} \max_{p \in \mathbf{P}} \mathbb{E}_{p,q}(r(T_p, I_q)) = \max_{p \in \mathbf{P}} \min_{q \in \mathbf{Q}} \mathbb{E}_{p,q}(r(T_p, I_q))$$

and

$$\min_{q \in \mathbf{Q}} \max_{T \in \mathcal{T}} \mathbb{E}_q(r(T, I_q)) = \max_{p \in \mathbf{P}} \min_{I \in \mathcal{I}} \mathbb{E}_p(r(T_p, I)).$$

As a consequence, for any $q \in \mathbf{Q}$ and $p \in \mathbf{P}$, we have

$$\max_{T \in \mathcal{T}} \mathbb{E}_q(r(T, I_q)) \geq \min_{I \in \mathcal{I}} \mathbb{E}_p(r(T_p, I)).$$

This inequality states that an upper bound on the inverse of the approximation ratio of the best universally truthful and Pareto optimal mechanism T_p on the worst instance is upper bounded by the inverse of the approximation ratio of the best deterministic truthful and Pareto optimal mechanism on any random instances. Hence, in order to bound $\min_{I \in \mathcal{I}} \mathbb{E}_p(r(T_p, I))$, we only need to construct an appropriate random instance and show the upper bound of the best deterministic truthful and Pareto optimal mechanism on this random instance. Consider the *triangle instance* where $N = \{1, 2, \dots, n_1\}$ and $A = \{a_1, a_2, \dots, a_{n_1}\}$, and an agent i ’s preference ordering is $a_1 \succ_i a_2 \succ_i \dots \succ_i a_i$, for any $i \in N$.

Let \mathcal{S} be the set of all the permutations of agents’ preference lists of the triangle instance. Consider now a random instance S_{uni} as the uniform distribution of \mathcal{S} . It is obvious that the output of any serial dictatorship mechanism (which is deterministic truthful and Pareto optimal) running on \mathcal{S} is the same. Hence, for any serial dictatorship mechanism (SDM), $\mathbb{E}_{uni}(r(SDM, S_{uni}))$ is equal to the inverse of the approximation ratio of RSDM (which is just SDM with uniformly random order of agents) when running on the triangle instance.

Online lower bound. We now apply these preliminaries to the online version of our problem. Recall that applicants in this online problem are truthful due to truthfulness of serial dictatorship mechanism. The strategy of the administrator is a random order in which the applicants are interviewed. More precisely, let Π denote the set of all the permutations of applicants and P be the set of probability distributions on Π . Let Π_p be a random order of applicants, where the order is selected according to the distribution $p \in P$ on Π , and then the strategy set of the administrator is $\{\Pi_p, p \in P\}$. We will show that the best strategy for the administrator is to select applicants’ order uniformly at random.

Theorem 6.1. *The best strategy for the administrator in the online problem is to select applicants’ order uniformly at random. Thus, any other randomized strategy, than the one used in Algorithm 3, would lead to an approximation guarantee worse than $\frac{e}{e-1}$.*

Proof. Let $\mathbb{E}_{p,q}(r(\Pi_p, I_q))$ be the inverse of the approximation ratio when the random order is Π_p and the random instance is I_q , and Π_{uni} denotes the uniform order. By the approximation ratio of RSDM, for any I , $\mathbb{E}_{uni}(r(\Pi_{uni}, I)) \geq \frac{e-1}{e}$. Now for upper bound of $\mathbb{E}_{p,q}(r(\Pi_p, I_q))$, by Yao's principle, $\max_{T \in \Pi} \mathbb{E}_q(r(T, I_q)) \geq \min_{I \in \mathcal{I}} \mathbb{E}_p(r(T_p, I))$. Recall that S_{uni} is the uniform distribution over \mathcal{S} . Then we need to upper bound $\max_{T \in \Pi} \mathbb{E}_q(r(T, S_{uni}))$, which in fact is equal to the inverse of the approximation ratio obtained by running RSDM on the triangle instance, which is $\frac{e-1}{e}$. The argument is as follows. Suppose object a_k is allocated by RSDM with probability $p_k \leq 1$ on the triangle instance. Then, because there are $n_1 - k + 1$ agents with a_k in their preference lists, each such agent obtains a_k with equal probability $\frac{p_k}{n_1 - k + 1}$. Therefore, agent i is allocated an object with probability $\sum_{j=1}^i \frac{p_j}{n_1 - j + 1}$, which is at most $\min\{1, \sum_{j=1}^i \frac{1}{n_1 - j + 1}\}$. Now, summing over all the agents, by a simple calculation we get that the expected cardinality of allocated agents is at most $n_1(1 - \frac{1}{e})$, for large enough n_1 . Hence, the approximation ratio is tight. \square

Lower bound for randomized mechanisms. If we can prove that the output of any deterministic truthful and Pareto optimal mechanism running on \mathcal{S} is the same as that of a serial dictatorship mechanism, then $\max_{T \in \mathcal{T}} \mathbb{E}_q(r(T, I)) = 1 - \frac{1}{e}$. To show our lower bound it suffices to show that the sum of the sizes of all the matchings output by any deterministic truthful and Pareto optimal mechanism that runs on \mathcal{S} is smaller than that of any serial dictatorship run on \mathcal{S} . Then $\max_{T \in \mathcal{T}} \mathbb{E}_q(r(T, I)) = 1 - \frac{1}{e}$. We use $\#\phi(\mathcal{S})$ to denote the sum of sizes of all the matchings output by mechanism ϕ when run on \mathcal{S} . Recall that SDM denotes a serial dictatorship, then we would like to prove that $\#\phi(\mathcal{S}) \leq \#^{SDM}(\mathcal{S})$, for any n_1 and n_2 and for any universally truthful and Pareto optimal mechanism ϕ . Until now, we can prove this inequality assuming $n_1 = n_2 = 3$, which gives the lower bound of $\frac{18}{13}$ for any universally truthful and Pareto optimal mechanism.

Theorem 6.2. *For any deterministic truthful and Pareto optimal mechanism ϕ , $\#\phi(\mathcal{S}) \leq 13$, when $n_1 = 3$. Thus, any universally truthful and Pareto optimal mechanism for this problem has an approximation ratio of at least $\frac{18}{13}$.*

Proof. Suppose the agents are 1, 2, 3 and objects are a, b, c . We use the notation $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & \\ a & b & \underline{c} \end{pmatrix}$ to denote assignments that allocate a to agent 1, b to agent 2 and c to agent

3, where row i denotes agent i 's preference list and preference ordering is the increasing order of column indices, $i = 1, 2, 3$. If there are no underlines of the objects, then this

notation denotes the input of mechanism. Note that in this setting, $\mathcal{S} = \left\{ \begin{pmatrix} a & & \\ a & b & \\ a & b & c \end{pmatrix}, \begin{pmatrix} a & b & \\ a & & \\ a & b & c \end{pmatrix}, \begin{pmatrix} a & b & c \\ a & & \\ a & b & \end{pmatrix}, \begin{pmatrix} a & b & c \\ a & b & \\ a & & \end{pmatrix}, \begin{pmatrix} a & b & c \\ a & & c \\ a & b & \end{pmatrix} \right\}$. We would

like to show that for any deterministic truthful and Pareto optimal mechanism ϕ , $\#\phi(\mathcal{S}) \leq 13$. W.l.o.g. suppose $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & c \\ a & b & \underline{c} \end{pmatrix}$, and we will consider the following two cases:

Case (i): If $\begin{pmatrix} a & b & \underline{c} \\ a & \underline{b} & c \\ \underline{a} & b & c \end{pmatrix}$, then we will show that $\begin{pmatrix} \underline{a} & b & \\ a & \underline{b} & c \\ a & b & \underline{c} \end{pmatrix}$. (Observe that the first agent must get a because otherwise we have contradiction with truthfulness by

$$\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & c \\ a & b & \underline{c} \end{pmatrix}.)$$

Now, if $\begin{pmatrix} \underline{a} & b \\ a & \underline{b} & c \\ a & b & \underline{c} \end{pmatrix}$ would not hold then $\begin{pmatrix} \underline{a} & b \\ a & b & \underline{c} \\ a & \underline{b} & c \end{pmatrix}$. Then we obtain $\begin{pmatrix} a & b \\ a & b \\ \underline{a} & b & c \end{pmatrix}$.

The reason is as follows: $\begin{pmatrix} a & b & \underline{c} \\ a & \underline{b} \\ \underline{a} & b & c \end{pmatrix}$ implies that the first agent in the input $\begin{pmatrix} a & b \\ a & b \\ a & b & c \end{pmatrix}$

cannot get any object by truthfulness. Similarly, from $\begin{pmatrix} \underline{a} & b \\ a & b & \underline{c} \\ a & \underline{b} & c \end{pmatrix}$, the second agent in

$\begin{pmatrix} a & b \\ a & b \\ a & b & c \end{pmatrix}$ cannot get any object by truthfulness. Thus we have that $\begin{pmatrix} a & b \\ a & b \\ \underline{a} & b & c \end{pmatrix}$,

which is a contradiction to Pareto optimality.

By a similar argument, we have $\begin{pmatrix} a & b \\ a & \underline{b} \\ \underline{a} & b & c \end{pmatrix}$, $\begin{pmatrix} \underline{a} & b \\ a & \underline{b} & c \\ a & b & \underline{c} \end{pmatrix}$ and $\begin{pmatrix} a \\ a & \underline{b} \\ \underline{a} & b & c \end{pmatrix}$. From

$\begin{pmatrix} a & b \\ a & \underline{b} \\ \underline{a} & b & c \end{pmatrix}$, we know the size of the matching output from $\begin{pmatrix} a & b \\ a \\ a & b & c \end{pmatrix}$ is 2. From

$\begin{pmatrix} a & b & \underline{c} \\ a & \underline{b} \\ \underline{a} & b & c \end{pmatrix}$, we know the size of the matching output from $\begin{pmatrix} a & b & c \\ a \\ a & b \end{pmatrix}$ is 2. From

$\begin{pmatrix} \underline{a} & b \\ a & \underline{b} & c \\ a & b & \underline{c} \end{pmatrix}$, we know the size of the matching output from $\begin{pmatrix} a & b \\ a & b & c \\ a \end{pmatrix}$ is 2. From

$\begin{pmatrix} \underline{a} \\ a & \underline{b} & c \\ a & b & \underline{c} \end{pmatrix}$, we know the size of the matching output from $\begin{pmatrix} a \\ a & b & c \\ a & b \end{pmatrix}$ is at most 2.

Thus, if the current mechanism is ϕ^1 then $\#\phi^1(\mathcal{S}) \leq 13$.

Case (ii): If $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} \\ a & b & \underline{c} \end{pmatrix}$, we consider the following two cases:

Case (ii-a): If $\begin{pmatrix} \underline{a} & b \\ a & b & \underline{c} \\ a & \underline{b} & c \end{pmatrix}$, then $\begin{pmatrix} \underline{a} & b \\ a & b \\ a & \underline{b} & c \end{pmatrix}$, and we conclude that $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & c \\ a & b \end{pmatrix}$,

otherwise suppose $\begin{pmatrix} a & \underline{b} & c \\ \underline{a} & b & c \\ a & b \end{pmatrix}$ (since $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & c \\ a & b \end{pmatrix}$), then b is allocated to agent 1 in

$\begin{pmatrix} a & b \\ a & b & c \\ a & b \end{pmatrix}$. From $\begin{pmatrix} \underline{a} & b \\ a & b & \underline{c} \\ a & \underline{b} & c \end{pmatrix}$, we know b is allocated to agent 3 in $\begin{pmatrix} a & b \\ a & b & c \\ a & b \end{pmatrix}$,

a contradiction. Hence, $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & c \\ a & b \end{pmatrix}$, then we know $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} \\ a & b \end{pmatrix}$ (from $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & c \\ a & b \end{pmatrix}$)

and $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} \\ a & b & \underline{c} \end{pmatrix}$, and $\begin{pmatrix} \underline{a} & b \\ a & b & \underline{c} \\ a & \underline{b} \end{pmatrix}$ (From $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & c \\ a & b \end{pmatrix}$ and $\begin{pmatrix} \underline{a} & b \\ a & b & \underline{c} \\ a & \underline{b} \end{pmatrix}$). Now the

matching size of assignment of $\begin{pmatrix} a & b & c \\ a & & \\ a & b & \end{pmatrix}$ and $\begin{pmatrix} a & b & c \\ a & b & \\ a & & \end{pmatrix}$ is both 2 (from $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & \\ a & b & \end{pmatrix}$).

The matching size of assignment of $\begin{pmatrix} a & b & c \\ a & b & c \\ a & & \end{pmatrix}$ is 2 since $\begin{pmatrix} \underline{a} & b & \\ a & b & \underline{c} \\ a & \underline{b} & \end{pmatrix}$. The matching

size of assignment of $\begin{pmatrix} a & b & \\ a & & \\ a & b & c \end{pmatrix}$ is 2 following from $\begin{pmatrix} \underline{a} & b & \\ a & b & \\ a & \underline{b} & c \end{pmatrix}$. Consider the assign-

ment of $\begin{pmatrix} a & & \\ a & b & c \\ a & b & c \end{pmatrix}$, no matter what the assignment is, at most one matching size of

assignment of $\begin{pmatrix} a & & \\ a & b & c \\ a & b & c \end{pmatrix}$ and $\begin{pmatrix} a & & \\ a & b & c \\ a & b & \end{pmatrix}$ is 3. Denote the mechanism in this case by

ϕ_2 , then $\#\phi_2(\mathcal{S}) \leq 13$.

Case (ii-b): If $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & c \\ a & b & \underline{c} \end{pmatrix}$, recall that $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & \\ a & b & \underline{c} \end{pmatrix}$ and $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & c \\ a & b & \underline{c} \end{pmatrix}$, consider the following two cases:

Case (ii-b-1): If $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & c \\ a & b & \end{pmatrix}$, then $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & \\ a & b & \end{pmatrix}$ since $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & c \\ a & b & \underline{c} \end{pmatrix}$. We know the

matching sizes of assignment of $\begin{pmatrix} a & b & c \\ a & b & c \\ a & & \end{pmatrix}$ and $\begin{pmatrix} a & b & c \\ a & & \\ a & b & \end{pmatrix}$ are both 2. Since $\begin{pmatrix} \underline{a} & b & \\ a & \underline{b} & \\ a & b & \underline{c} \end{pmatrix}$

due to $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & \\ a & b & \underline{c} \end{pmatrix}$ and $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & c \\ a & b & \end{pmatrix}$, the matching size of assignment of $\begin{pmatrix} a & b & \\ a & & \\ a & b & c \end{pmatrix}$

is 2. Since $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & c \\ a & b & \end{pmatrix}$ due to $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & c \\ a & b & \underline{c} \end{pmatrix}$ and $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & c \\ a & b & \underline{c} \end{pmatrix}$, then the matching size

of assignment of $\begin{pmatrix} a & b & c \\ a & b & c \\ a & & \end{pmatrix}$ is 2. Similar as the above argument, consider the assign-

ment of $\begin{pmatrix} a & & \\ a & b & c \\ a & b & c \end{pmatrix}$, no matter what the assignment is, at most one matching size of

assignment of $\begin{pmatrix} a & & \\ a & b & c \\ a & b & c \end{pmatrix}$ and $\begin{pmatrix} a & & \\ a & b & c \\ a & b & \end{pmatrix}$ is 3. Denote the mechanism in this case by

ϕ_3 , then $\#\phi_3(\mathcal{S}) \leq 13$.

Case (ii-b-2): If $\begin{pmatrix} a & \underline{b} & c \\ \underline{a} & b & c \\ a & b & \end{pmatrix}$, recall that we have $\begin{pmatrix} \underline{a} & b & \\ a & \underline{b} & c \\ a & b & \underline{c} \end{pmatrix}$, $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & \\ a & b & \underline{c} \end{pmatrix}$ and

$\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & c \\ a & b & \underline{c} \end{pmatrix}$. From $\begin{pmatrix} a & \underline{b} & c \\ \underline{a} & b & c \\ a & b & \end{pmatrix}$ and $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & c \\ a & b & \underline{c} \end{pmatrix}$, we get $\begin{pmatrix} a & \underline{b} & \\ \underline{a} & b & c \\ a & b & \end{pmatrix}$, then the

matching sizes of assignment of $\begin{pmatrix} a & & \\ a & b & c \\ a & b & \end{pmatrix}$ and $\begin{pmatrix} a & b & \\ a & b & c \\ a & & \end{pmatrix}$ are both 2. From $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & \\ a & b & \underline{c} \end{pmatrix}$ and $\begin{pmatrix} \underline{a} & b & \\ a & \underline{b} & c \\ a & b & \underline{c} \end{pmatrix}$, we get $\begin{pmatrix} \underline{a} & b & \\ a & \underline{b} & \\ a & b & \underline{c} \end{pmatrix}$, then the matching size of assignment of $\begin{pmatrix} a & b & \\ a & & \\ a & b & c \end{pmatrix}$ is 2. From $\begin{pmatrix} \underline{a} & b & c \\ a & \underline{b} & \\ a & b & \underline{c} \end{pmatrix}$ and $\begin{pmatrix} a & \underline{b} & c \\ \underline{a} & b & \\ a & b & \end{pmatrix}$, we get $\begin{pmatrix} a & \underline{b} & c \\ \underline{a} & b & \\ a & b & \end{pmatrix}$, then the matching size of assignment of $\begin{pmatrix} a & b & c \\ a & b & \\ a & & \end{pmatrix}$ is 2. From $\begin{pmatrix} a & \underline{b} & c \\ \underline{a} & b & \\ a & b & \end{pmatrix}$ and $\begin{pmatrix} \underline{a} & b & \\ a & \underline{b} & c \\ a & b & \underline{c} \end{pmatrix}$, it follows that $\begin{pmatrix} a & \underline{b} & \\ \underline{a} & b & \\ a & b & \end{pmatrix}$, we conclude $\begin{pmatrix} \underline{a} & & \\ a & \underline{b} & \\ a & b & \underline{c} \end{pmatrix}$ is not true. Otherwise from $\begin{pmatrix} \underline{a} & & \\ a & \underline{b} & \\ a & b & \underline{c} \end{pmatrix}$ and $\begin{pmatrix} a & \underline{b} & \\ \underline{a} & b & \\ a & b & \end{pmatrix}$, it follows that $\begin{pmatrix} a & & \\ \underline{a} & b & \\ a & b & \end{pmatrix}$, which contradicts to the Pareto optimality of

the mechanism. Hence, the matching size of assignment of $\begin{pmatrix} a & & \\ a & b & \\ a & b & c \end{pmatrix}$ is 2. It is obvious to see that the matching size of assignment of $\begin{pmatrix} a & b & c \\ a & & \\ a & b & \end{pmatrix}$ is at most 3. Denote the

current mechanism as ϕ^4 , we know that $\#\phi^4(\mathcal{S}) \leq 13$. \square

Note that Theorem 6.2 shows that $\min_{I \in \mathcal{I}} \mathbb{E}_p(r(T_p, I)) \leq \max_{T \in \mathcal{T}} \mathbb{E}_q(r(T, S_{uni})) \leq \frac{13}{18}$, for any $p \in \mathcal{P}$, and S_{uni} is the uniform distribution over \mathcal{S} . Hence, the approximation ratio is at least $\frac{18}{13}$.

Lower bound for non-bossy mechanisms. We first briefly define the concept of non-bossiness for a deterministic mechanism (see, e.g., [23]). A deterministic mechanism ϕ is non bossy, if for any strict preference list $L(i)$, $L'(i)$, $L(-i)$ and $i \in N$, if $\phi_i(L(i), L(-i)) = \phi_i(L'(i), L(-i))$ then $\phi(L(i), L(-i)) = \phi(L'(i), L(-i))$. Pápai [23] proved that non-bossiness with truthfulness is equivalent to group truthfulness. Pycia and Ünver [26] characterized all the mechanisms which are group strategy-proof (note, strategy-proofness is a synonym with truthfulness) and Pareto optimal by trading cycle with owners and brokers mechanisms. Bade [7] showed that any mechanism which is group strategy-proof and Pareto optimal is r -equivalent to serial dictatorship, in the sense that if the order of agents is generated uniformly at random, the resulting random matching from serial dictatorship mechanism is the same as that from any group strategy-proof and Pareto optimal mechanism. Hence, by arguments in the preliminaries that use Yao's principle, we have the following tight lower bound for any group strategy-proof and Pareto optimal mechanism.

Theorem 6.3. *No universally truthful, non-bossy and Pareto optimal mechanism can achieve the approximation ratio better than $\frac{e}{e-1}$.*

Note that our mechanism with strict preference list and weights is universally truthful, nonbossy and Pareto optimal.

7 Conclusion

Whilst this paper has focused on Pareto optimality in the HA context, stronger forms of optimality are possible. For example, *minimum cost* (or *maximum utility*), *rank-maximal* and *popular* matchings can also be studied in the HA context, and a matching of each of these types is Pareto optimal (see, e.g., [22, Sec. 1.5] for definitions). As Pareto optimality is a unifying feature of all of these other forms of optimality, we chose to concentrate on this concept in our search for randomized truthful mechanisms that can provide good approximations to maximum matchings with desirable properties. Note that the lower bound on the performance of deterministic truthful mechanisms that produce Pareto optimal matchings extends to those producing matchings that satisfy these stronger optimality criteria. It will thus be the focus of future work to consider the performance of randomized truthful mechanisms for these problems.

Acknowledgements

We would like to thank Anna Bogomolnaia and Hervé Moulin for helpful discussions concerning the results in this paper. We would also like to thank anonymous reviewers for their valuable remarks on this paper.

References

- [1] A. Abdulkadiroğlu and T. Sönmez. Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica*, 66(3):689–701, 1998.
- [2] D.J. Abraham, K. Cechlárová, D.F. Manlove, and K. Mehlhorn. Pareto optimality in house allocation problems. In *Proc. ISAAC '04*, volume 3341 of *LNCS*, pages 3–15. Springer, 2004.
- [3] G. Aggarwal, G. Goel, C. Karande, and A. Mehta. Online vertex-weighted bipartite matching and single-bid budgeted allocations. In *Proc. SODA '11*, pages 1253–1264. ACM-SIAM, 2011.
- [4] H. Aziz, F. Brandt, and P. Harrenstein. Pareto optimality in coalition formation. *Games and Economic Behavior*, 82:562 – 581, 2013.
- [5] H. Aziz, S. Gaspers, S. Mackenzie, and T. Walsh. Fair assignment of indivisible objects under ordinal preferences. In *Proc. AAMAS 2014*, to appear, May 2014.
- [6] M. Babaioff, N. Immorlica, D. Kempe, and R. Kleinberg. Online auctions and generalized secretary problems. *SIGecom Exchanges*, 7(2), 2008.
- [7] S. Bade. Random serial dictatorship: The one and the only. Unpublished manuscript, available from <http://www.coll.mpg.de/sites/www.coll.mpg.de/files/workshop/newproof5.pdf>, 2013.
- [8] A. Bhalgat, D. Chakrabarty, and S. Khanna. Social welfare in one-sided matching markets without money. In *Proc. APPROX+RANDOM '11*, volume 6845 of *LNCS*, pages 87–98. Springer, 2011.
- [9] A. Bogomolnaia and H. Moulin. A new solution to the random assignment problem. *Journal of Economic Theory*, 100(2):295 – 328, 2001.

- [10] A. Bogomolnaia and H. Moulin. Random matching under dichotomous preferences. *Econometrica*, 72(1):257–279, 2004.
- [11] A. Bogomolnaia and H. Moulin. Size versus fairness in the assignment problem. Unpublished manuscript, available from http://www.gla.ac.uk/media/media_295542_en.pdf, 2013.
- [12] K. Cechlárová, P. Eirínakis, T. Fleiner, D. Magos, I. Mourtos, and E. Potpinková. Pareto optimality in many-to-many matching problems. Technical report, IM Preprint series A, no. 4/2013, PJ Šafárik University, Faculty of Science, Institute of Mathematics, December 2013.
- [13] D. Chakrabarty and C. Swamy. Welfare maximization and truthfulness in mechanism design with ordinal preferences. In *Proc. ITCS 2014*, pages 105–120. ACM, May 2014.
- [14] N. Chen and A. Ghosh. Algorithms for Pareto stable assignment. In V. Conitzer and J. Rothe, editors, *Proc. COMSOC '10*, pages 343–354. Düsseldorf University Press, 2010.
- [15] N. R. Devanur, K. Jain, and R. Kleinberg. Randomized primal-dual analysis of RANKING for online bipartite matching. In *Proc. SODA '13*, pages 101–107. ACM-SIAM, 2013.
- [16] N.B. Dimitrov and C.G. Plaxton. Competitive weighted matching in transversal matroids. *Algorithmica*, 62(1-2):333–348, 2012.
- [17] S. Dughmi and A. Ghosh. Truthful assignment without money. In *Proc. EC '10*, pages 325–334. ACM, 2010.
- [18] P. Gärdenfors. Assignment problem based on ordinal preferences. *Management Science*, 20(3):331–340, 1973.
- [19] A. Hylland and R. Zeckhauser. The efficient allocation of individuals to positions. *Journal of Political Economy*, 87(2):293–314, 1979.
- [20] P. Jaramillo and V. Manjunath. The difference indifference makes in strategy-proof allocation of objects. *Journal of Economic Theory*, 147(5):1913–1946, 2012.
- [21] R. M. Karp, U. V. Vazirani, and V. V. Vazirani. An optimal algorithm for on-line bipartite matching. In *Proceedings of the Twenty-second Annual ACM Symposium on Theory of Computing*, STOC '90, pages 352–358. ACM, 1990.
- [22] D.F. Manlove. *Algorithmics of Matching Under Preferences*. World Scientific, 2013.
- [23] S. Pápai. Strategyproof assignment by hierarchical exchange. *Econometrica*, 68(6):1403–1433, 2000.
- [24] C. G. Plaxton. A simple family of top trading cycles mechanisms for housing markets with indifference. Unpublished manuscript, available from <http://www.cs.utexas.edu/~plaxton/pubs/2013/icgt.pdf>, 2013.
- [25] A.D. Procaccia and M. Tennenholtz. Approximate mechanism design without money. *ACM Transactions on Economics and Computation*, 1(4), 2013. Article number 18.
- [26] M. Pycia and M.U. Ünver. Incentive compatible allocation and exchange of discrete resources. Unpublished manuscript, available from <http://pycia.bol.ucla.edu/pycia-unver-tcbo.pdf>, 2014.

- [27] A.E. Roth and A. Postlewaite. Weak versus strong domination in a market with indivisible goods. *Journal of Mathematical Economics*, 4:131–137, 1977.
- [28] A.E. Roth, T. Sonmez, and M.U. Ünver. Pairwise kidney exchange. *Journal of Economic Theory*, 125(2):151–188, December 2005.
- [29] D. Saban and J. Sethuraman. The complexity of computing the random priority allocation matrix. In *Proc. WINE '13*, volume 8289 of *LNCS*, page 421. Springer, 2013. Full version available at <http://www.columbia.edu/~js1353/pubs/rpcomplexity.pdf>.
- [30] D. Saban and J. Sethuraman. House allocation with indifference: A generalization and a unified view. In *Proc. EC '13*, pages 803–820. ACM, 2013.
- [31] L.-G. Svensson. Queue allocation of indivisible goods. *Social Choice and Welfare*, 11(4):323–330, 1994.
- [32] T. Uno. Algorithms for enumerating all perfect, maximum and maximal matchings in bipartite graphs. In *Proc. ISAAC '97*, volume 1350, pages 92–101. Springer, 1997.
- [33] A.C.-C. Yao. Probabilistic computations: Toward a unified measure of complexity (extended abstract). In *Proc. FOCS '77*, pages 222–227. IEEE Computer Society, 1977.
- [34] L. Zhou. On a conjecture by Gale about one-sided matching problems. *Journal of Economic Theory*, 52(1):123–135, 1990.